Disordered binding of small molecules to a (12-28)

Convertino, M ; Vitalis, A ; Caflisch, A

Abstract: In recent years, an increasing number of small molecules and short peptides have been identified that interfere with aggregation and/or oligomerization of the Alzheimer -amyloid peptide (A). Many of them possess aromatic moieties, suggesting a dominant role for those in interacting with A along various stages of the aggregation process. In this study, we attempt to elucidate whether interactions of such aromatic inhibitors with monomeric A (12-28) point to a common mechanism of action by performing atomistic molecular dynamics simulations at equilibrium. Our results suggest that, independently of the presence of inhibitors, monomeric A (12-28) populates a partially collapsed ensemble that is largely devoid of canonical secondary structure at 300 K and neutral pH. The small molecules have different affinities for A (12-28) that can be partially rationalized by the balance of aromatic and charged moieties constituting the molecules. There are no predominant binding modes, although aggregation inhibitors preferentially interact with the N-terminal portion of the fragment (residues 13-20). Analysis of the free energy landscape of A (12-28) reveals differences highlighted by altered populations of a looplike conformer in the presence of inhibitors. We conclude that intrinsic disorder of A persists at the level of binding small molecules and that inhibitors can significantly alter properties of monomeric A via multiple routes of differing specificity.

DOI: https://doi.org/10.1074/jbc.M111.285957

JBC Papers in Press. Published on October 3, 2011 as Manuscript M111.285957
The latest version is at http://www.jbc.org/cgi/doi/10.1074/jbc.M111.285957

DISORDERED BINDING OF SMALL MOLECULES TO Aβ12-28*

Marino Convertino, Andreas Vitalis, and Amedeo Caflisch

Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

*Running head: *Disordered binding of small molecules to Aβ12-28*

Address correspondence to:
Prof. Dr. Amedeo Caflisch, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland.
Tel: +41 44 635 55 21; fax: +41 44 635 68 62; E-mail: caflisch@bioc.uzh.ch.
Dr. Andreas Vitalis, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland.
Tel: +41 44 635 55 97; fax: +41 44 635 68 62; E-mail: a.vitalis@bioc.uzh.ch.

Keywords: Alzheimer's disease; intrinsically disordered proteins; molecular dynamics; protein aggregation; aggregation inhibitors

Background: Inhibition of pathogenic protein aggregation by small molecules is poorly understood.

Results: Aggregation inhibitors alter the free energy landscape of a relevant fragment of Alzheimer's amyloid-β protein in subtle but complex fashion.

Conclusion: Intrinsic disorder of disease proteins persists at the level of binding small molecules.

Significance: Hallmark characteristics and efficacy of inhibitors can be reconciled with lack of specificity.

SUMMARY

In recent years, an increasing number of small molecules and short peptides have been identified that interfere with aggregation and/or oligomerization of the Alzheimer's β-amyloid peptide (Aβ). Many of them possess aromatic moieties suggesting a dominant role for those in interacting with Aβ along various stages of the aggregation process. In this study, we attempt to elucidate whether interactions of such aromatic inhibitors with monomeric Aβ12-28 point to a common mechanism of action by performing atomistic molecular dynamics simulations at equilibrium. Our results suggest that, independently of the presence of inhibitors, monomeric Aβ12-28 populates a partially collapsed ensemble that is largely devoid of canonical secondary structure at 300 K and neutral pH. The small molecules have different affinities for Aβ12-28 that can be partially rationalized by the balance of aromatic and charged moieties constituting the molecules. There are no predominant binding modes, even though aggregation inhibitors preferentially interact with the N-terminal portion of the fragment (residues 13-20). Analysis of the free energy landscape of Aβ12-28 reveals differences highlighted by altered populations of a loop-like conformer in the presence of inhibitors. We conclude that intrinsic disorder of Aβ persists at the level of binding small molecules, and that inhibitors can significantly alter properties of monomeric Aβ via multiple routes of differing specificity.

Alzheimer's disease (AD) is the most common form of dementia in the elderly. Strong genetic, physiological and biochemical evidence suggests that the β-amyloid peptide (Aβ) plays a key role in the pathogenesis of AD (1). Neuropathological changes in the brain of AD patients include neuronal death in the regions related to memory and cognition, as well as the abnormal presence of intra- and extracellular protein aggregates (2,3) known as neurofibrillary tangles and amyloid plaques. They are the final results of a complex series of oligomerization and polymerization events that typically follow a nucleation-dependent mechanism (4). The left-hand side of

Fig. 1 shows a cartoon schematic illustrating a few steps along the pathway of Aβ fibrillization. The nucleus is typically assumed to be a larger oligomer (4-6), and the nucleation event itself may be linked to a critical structural transition involving tertiary and quaternary contacts within such an oligomer or protofibril (5). Subsequent monomer addition appears to be the dominant mode of fibril elongation (4). Peptide aggregation processes have been studied in depth with several experimental (7,8) and computational techniques (9-12), but often remain poorly understood. Although little is known on the link between the aggregation mechanism and neurotoxicity (13), experimental evidence indicates that soluble oligomers and fibrillar precursors of Aβ may be the dominant neurotoxic species (14).

In recent years, increasing evidence points to a link between disease and disorder, specifically, the functions and properties of intrinsically disordered proteins (IDPs) and polypeptide stretches within proteins (IDRs) (15,16). The ensembles explored by such sequences, which are estimated to make up about 20% (17) of eukaryotic genomes, are highly diverse and devoid of long-lived, "folded" conformers (18). Extensive analyses have shown that simple sequence-based classifiers such as mean hydrophobicity or net charge can be used to distinguish folded proteins from IDPs (19). $A\beta_{40/42}$ belongs to the class of collapsed-disordered IDPs (20) on account of its low net charge and high hydrophobicity (21,22). IDPs often attain partial order upon functional or deleterious interactions with folded proteins or with other IDPs (23). Indeed, pathogenic self-assembly can be viewed as a specific variant of the latter case. Given that collapse and aggregation are guided by the same driving forces, it is perhaps not surprising that IDPs such as Aβ or polyglutamine are associated with protein aggregation diseases (24-26).

Inherently, structural drug design aimed at finding compounds that interfere with an IDP-mediated process faces the challenge that structural targets emerge only later on the pathway. Nevertheless, the identification and detailed biophysical characterization of small molecules that modulate Aβ peptide self-assembly are expected to generate new lead candidates for clinical studies. Several therapeutic strategies have been suggested for blocking key-steps in the amyloid aggregation process, including the direct inhibition of aggregation by using either peptides or small molecules (27-38). As an example, indole derivatives inhibited fibril formation of Aβ peptide (39,40) and lysozyme (41). Anthraquinones were shown to be inhibitors of tau protein (42) and $A\beta_{40}$ aggregation (37), and hybrid molecules bearing both indole and quinone rings have been effective in the recovery of a fly model of AD (43). In addition, antioxidants, e.g., resveratrol (44,45) and epigallocatechin-3-gallate (46), and non-steroidal anti-inflammatory molecules such as naproxen (47,48) revealed new biological activities in the inhibition of amyloid aggregation.

Recent X-ray microcrystallography (49,50) and solid-state nuclear magnetic resonance (NMR) spectroscopy (51) studies have provided atomistic information on the interactions between small-molecule binders and amyloid fibrils. Fig. 1 illustrates why this may be less relevant than the interactions of inhibitors with soluble peptide species. In essence, compounds that specifically bind fibrils (step VI in Fig. 1) may destabilize the latter (52), but will have little impact on the association and conformational equilibria prior to nucleation (steps I and II). An alternative mechanism of inhibition could be a depletion of nucleation-competent (53) and/or toxic oligomers (54) either by stabilization of low molecular weight species such as monomers and dimers (steps III and IV) or of larger off-pathway oligomers (step V). Given the disordered nature of the binding partner (55), it is quite reasonable to stipulate that small molecules can have differential effects for all indicated steps including the known ability to increase the rate of fibrillization (56,57). The complexity is exacerbated by the fact that the dominant pathways may shift in dose-dependent manner. Fig. 1 implies that studies of the binding equilibria of monomeric Aβ with inhibitors can yield insight regarding the molecular mechanisms of inhibition.

Here, we use molecular dynamics (MD) simulations to analyze ten different small-molecule inhibitors of Aβ peptide aggregation, and focus on their influence on the free energy surface of monomeric $A\beta_{12-28}$. The choice of studying a truncated construct is motivated by three reasons. First and foremost, it is important to be able to obtain statistically reliable simulation results.

Conformational transitions in the full length alloforms can occur on timescales that exceed the currently accessible regime (58). Second, residues 12-28 highlight the role of the so-called central hydrophobic cluster, CHC (59), residues 17-21, that is often assumed to be critical in mediating peptide-inhibitor (60) as well as peptide-peptide interactions (61). By discarding the hydrophobic C-terminus, our simulations allow us to delineate the possible specific roles played by the residues in this stretch. Truncation of residues 1-11 is justified by experimental (62) and simulation (21) studies that show this segment to be completely unstructured. Third, segment 12-28 could be used in high enough concentrations suitable for solution NMR spectroscopy experiments (43), because it has lower oligomerization and fibrillization propensities than Aβ40 and Aβ42. This implies that it will be easier to derive testable hypotheses. We use the cut-based free energy profile (cFEP) method (63) to identify the metastable states of monomeric Aβ12-28 and the change of their relative stability upon inhibitor binding.

This study was inspired by the following questions: How does the free energy surface of monomeric Aβ12-28 change in the presence of small molecules that are known to interfere with oligomerization and/or fibril formation? Do different inhibitors of Aβ peptide self-assembly share similar interaction motifs with monomeric Aβ12-28? Is there a major binding mode? The MD simulation results indicate that monomeric Aβ12-28 is largely disordered with and without inhibitors. The most frequent interaction motifs are similar for different inhibitors. There is no predominant binding mode because Aβ12-28 is highly flexible, and its plasticity is marginally influenced by the small-molecule inhibitors. An analysis of binding frequency and the enhancement of a specific, otherwise transiently populated conformation of Aβ12-28 in the presence of inhibitors, suggests a complex interplay of interfacial effects, trends that can be mapped back to simple physicochemical properties of the primary sequence, and lastly highly specific effects that require elucidation by atomistic simulations.

EXPERIMENTAL PROCEDURES

Implicit solvent simulations. Simulations were performed with the CHARMM program (64). The Aβ12-28 peptide and inhibitors were modeled using the united atoms CHARMM PARAM19 force field with its default truncation scheme for non-bonded interactions (cutoff of 7.5 Å). Parameters for 1,4-napthoquinon-2-yl-L-tryptophan (NQTrp), anthracene, and 9,10-anthraquinone were derived as reported in (37,43). Protonation states of titratable residues were considered at neutral pH. In particular, the His side chains of Aβ12-28 and β-Ala-His were neutral (protonated at the Nδ), while the charges of the Asp/Glu and Arg/Lys side chains were -1 and +1 electronic units, respectively. The net charge of the Aβ12-28 segment is zero as there are two positively charged residues (Lys16 and Lys28) and two negatively charged residues (Glu22 and Asp23), and the N- and C-terminus were capped with acetyl and N-methylamide groups, respectively. The electrostatic contribution to solvation was accounted for by using FACTS (65), an efficient generalized Born implicit solvent model based on the fully analytical evaluation of the volume and spatial symmetry of the solvent that is displaced from around a solute atom by its neighboring atoms. The non-polar contribution to the total effective solvation energy was approximated by a term proportional to the solvent-accessible surface area of the solute using a surface tension-like, multiplicative parameter of 7.5 cal mol^{-1} $Å^{-2}$. Starting conformations were prepared by placing fully extended Aβ12-28 in the presence or absence of a single molecule of the inhibitor in the simulation box (1:1 concentration ratio). Simulations were carried out with periodic boundary conditions at fixed peptide concentration of ca. 2.5 mM (87 Å cubic simulation box) using the Langevin integrator at low friction (coefficient of 0.15 ps^{-1}) and at a temperature of 300 K. Using a time step of 2 fs, for each system we performed three independent runs of 5 μs each.

Explicit solvent simulations. Using GROMACS v4.5.3 (66), capped Aβ12-28 was simulated in a cubic box of 60 Å side length in the NPT ensemble. The velocity rescaling thermostat of Bussi *et al*. (67) was used to keep a constant temperature of 310 K, while an ambient pressure of 1 bar was maintained using the Parrinello-Rahman barostat (68). The peptide was represented with the CHARMM27 all-atom force field including CMAP corrections (69). The bath

consisted of a solution of ca. 150 mM NaCl in TIP3P water (70). Electrostatic interactions were modeled by the particle-mesh Ewald method (71). All real-space interactions were truncated at 12 Å. Neighbor lists were recalculated every 5 steps. LINCS (72) was used to constrain all bonds involving hydrogen atoms to their parameter-derived values. The time step was 2 fs, and we obtained three independent simulations starting from random, extended structures that each are 380 ns in length, the first 20 ns of which we discarded as equilibration. Preliminary analyses of secondary structure propensities or contact patterns revealed that given the reduced amount of sampling and increased friction statistical convergence for the majority of readouts could not be obtained. Therefore, data from explicit solvent simulations are only included in Fig. 2.

cFEP analysis. The 750,000 snapshots of each system were clustered by the Leader algorithm as implemented in Wordom (73) using the Cα atoms of residues 14-24 and a threshold of 1.0 Å. The cut-based Free Energy Profile (cFEP) (63) technique was used to identify metastable states of monomeric Aβ12-28 and the change of their relative stability upon inhibitor binding. The input for the cFEP calculation is the network of conformational transitions, which is derived from the direct transitions between clusterized snapshots (nodes of the network) sampled at a given time interval (20 ps here) along the MD simulations. For each node, nodes are partitioned into two groups using the values of the mean first passage times (mfpt) to the reference node to define a cut. The free energy is related to the maximum flow across the cut and approximated as $\Delta G=-kT\ln(Z_{AB})$, where Z_{AB} is the partition function of the mfpt-based cutting surface (for further details refer to Ref. (63)). The result is a one-dimensional profile along a reaction coordinate (the relative partition function, termed Z_A/Z) that preserves the barrier height between the free energy basins.

RESULTS AND DISCUSSION

The following analysis concerns the FACTS implicit solvent MD simulations at 300 K. The reference system is monomeric Aβ12-28 while the simulations with inhibitors contained a single inhibitor molecule along with Aβ12-28 at a final concentration of ca. 2.5 mM. Each of these 11 systems was simulated for a total of 15 μs. If not stated otherwise, the statistical significance (*i.e.*, convergence) of the simulation results was assessed by computing min/max errors from block averages over three or six blocks.

Monomeric Aβ12-28 is partially collapsed and disordered. Visual inspection of the trajectories showed that Aβ12-28 does not attain any specific, long-lived structure akin to a folded ensemble. Conformational transitions are rapid and yield a disordered ensemble. To quantitatively assess the overall polymeric state of the Aβ peptide, we computed the scaling of internal distances with sequence spacing as well as the angular correlation function as described in previous work (74). The data in Fig. 2 indicate that - independently of the presence of an inhibitor - the peptide populates a partially collapsed ensemble. Inhibitors appear to be able to cause both compaction and swelling of Aβ12-28, but many of the differences are insignificant. For fully globular species one would expect the internal scaling plot to yield a plateau (75) similar to what we observed for the full length alloforms (21), while the theoretical prediction for a chain in a good solvent is significantly more expanded. As described in Experimental Procedures, we performed calculations in explicit solvent of the free system. These data are shown as well in Fig. 2. They crudely illustrate similarity between implicit and explicit solvent ensembles (Panel A), and simultaneously demonstrate the difficulty to obtain converged explicit solvent results even for low-dimensional readouts such as the angular correlation function (Panel B).

To demonstrate dominant disorder in Aβ12-28, we show in Fig. 3 the cut-based free energy profile (cFEP) of the free system. The profile is devoid of significant free-energy barriers, and dominated by a large entropic basin composed of conformations that lack canonical secondary structure. The remaining 30-40% of the ensemble is made up by various enthalpic basins, and cartoon representations of representative snapshots have been added for illustration purposes. Taken together, these results indicate that Aβ12-28 is unstructured at physiological conditions in the absence of small-molecule inhibitors of aggregation. Similar observations hold for Aβ12-28 in the presence of any of the aggregation inhibitors

(see Figs. S2-S11 in the Supplemental Data). The lack of a predominant structure and the small height of barriers for the ensemble of monomeric Aβ12-28 are consistent with data from NMR spectroscopy experiments (76,77).

Main interactions between monomeric Aβ12-28 and aggregation inhibitors. Table 1 shows estimates for how frequently each inhibitor is "bound" to the peptide in the simulations. To resolve which parts of Aβ12-28 inhibitors bind to, Fig. 4 shows intra- and intermolecular contact maps for the free system as well as for simulations in the presence of inhibitors. The most frequently observed intermolecular contacts are shared by all inhibitors, and involve mainly the N-terminal segment (residues 13-20). With the exception of β-Ala-His that largely remains dissociated, the large hydrophobic side chains of Phe19 and Phe20 constitute the site of highest interaction probability in all cases. It is interesting to note that the larger peptidic inhibitors appear to show more specific contact patterns than for example anthracene or 9,10-anthraquinone. This is consistent with the complexity of the interfaces presented by each molecule: anthracene is a completely homogeneous tricyclic molecule, whereas the tetrapeptides have three different aromatic moieties (two of which have polar sites), one or two formal charges, and a mainly rigid spacer. The segment 24-28 appears to be largely inert to inhibitor binding consistent with the fact that residues in this region do not feature prominently in sequence analyses (62,78,79), mutation (80), or fragment binding studies (61) as being responsible for amyloid formation. While this is also true for the N-terminal residues we truncated, the same cannot be said for C-terminal residues 29-40/42. *A priori*, we have no reason to assume that inhibitors would not bind to this portion of Aβ (see Fig. 1). However, the questions in this work center around the interaction of aromatic inhibitors with the central hydrophobic core as the suspected site of highest affinity (81). Given that several of the inhibitors have significant impact on oligomerization and fibrillization (36,43) even at sub-equimolar concentrations ratios, this may represent a reasonable simplification.

We can attempt to explain the interaction patterns within the confines of the employed computational model by decomposing system energetics. As an example, we focus on NQTrp, which has the highest affinity for Aβ12-28 in the MD simulations (see Table 1) and is also one of the most potent inhibitors of Aβ peptide aggregation in experiments (43). In Fig. 5, the decomposition of the dominant term of the intermolecular energy into contributions from individual pairs of functional groups shows that the two aromatic moieties of NQTrp make favorable van der Waals interactions with the entire N-terminal stretch of the peptide (residues 13-20). In addition, binding appears to feature favorable electrostatic interactions involving the (charged) carboxy group of NQTrp and peptide residues 12-16, which possess a wealth of polar hydrogens (His13, His14, Gln15, and Lys16). This can be inferred indirectly from the favorable van der Waals interactions between the carboxy group and the peptide. We did not consider electrostatic interactions here because of effective multi-body terms preventing pairwise decomposition. In general, good correspondence is seen between the interaction energy matrix and contact maps indicating that binding is largely enthalpic. This is expected in particular given that entropic contributions to ligand binding stemming from the solvent are accounted for implicitly in the continuum model.

It is possible to rationalize the relative affinities for Aβ12-28 to the ten inhibitors by focusing on the three main contributions, *i.e.*, unfavorable desolvation of hydrophilic moieties, favorable burial of hydrophobic moieties, and favorable van der Waals interactions (see Tables 1 and S2). The total polar desolvation penalty ranges from 1 to 5 kcal/mol and scales roughly with the size of the inhibitor. No persistent salt bridges are formed and the contribution is generally unfavorable. Burial of hydrophobic surface contributes favorably. This term correlates with the number of aromatic moieties present on the inhibitor. Lastly, dispersive interactions between inhibitor and Aβ12-28 contribute the bulk of the favorable effective binding energy. It appears as if this contribution becomes less predictable with increasing numbers of hydrophobic moieties present on the inhibitor (*e.g*, NQTrp has a significantly more favorable van der Waals contribution than either Tyr-Aib-Trp-Phe or Tyr-Pro-Trp-Phe despite possessing fewer aromatic moieties). The computed effective binding energies reproduce the rank order seen in

Table 1 reasonably well. Two exceptions shall be mentioned: The first is Tyr-Pro-Phe-Phe, which appears to incur a significant entropic penalty upon binding that is consistent with a significantly more structured free energy surface of Aβ_{12-28} in the presence of Tyr-Pro-Phe-Phe (see Fig. S7 in Supplemental Data). On the contrary, anthracene exhibits high affinity despite its modest effective binding energy. Here, the entropic penalty at the inhibitor level is essentially zero (rigid molecule), and the contact map indicates the most degenerate binding (see Fig. 4) suggesting a minimal entropic penalty at the peptide level as well.

Returning to Fig. 4, we note that inhibitor binding has an impact on intramolecular contacts of Aβ_{12-28} as well. While the contact maps continue to indicate degenerate long-range interactions within the peptide, the presence of inhibitors seems to enhance contacts between residues 12-15 and 21-26 (marked in the plots). This is a relatively weak effect that is observed for all inhibitors but β-Ala-His, and most prominently for NQTrp. It points to the ability of molecules composed of similar building blocks to exert a generic effect on the conformational properties of Aβ_{12-28}, and this is discussed next.

Changes in the free energy surface of Aβ_{12-28} upon inhibitor binding. The DSSP strings in Figs. S1-S11 in Supplemental Data suggest that the overall secondary structure content of Aβ_{12-28} is not strongly altered in the presence of inhibitors compared to the free system (Fig. 3). Fig. 6 shows that the inhibitors generally increase the content of loop, bend, and turn conformations in the segment 13-23 at the expense of regular secondary structure, particularly of helix. These changes are qualitatively similar for all inhibitors. As expected, the low affinity compounds β-Ala-His and ^{D}Trp-Aib exhibit only weak effects. The same is true for anthracene despite its affinity being the second highest (Table 1). This is consistent with the entropic binding mode described above. Over the same sequence, β-secondary structure is significantly enhanced only in the presence of ^{D}Trp-Aib and Tyr-Pro-Phe-Phe. In contrast to residues 13-23, there are negligible differences for residues 24-28, which is consistent with the lack of interactions between the inhibitors and the C-terminal segment of Aβ_{12-28} (see Fig. 4). The effects are overall subtle and confirm the prevalence of disorder in binding.

As noted above, the presence of inhibitors does seem to enhance intramolecular Aβ_{12-28} contacts between residues 12-15 and 21-26 (see Fig. 4). This points to the potential of inhibitors to alter the free energy surface of Aβ_{12-28} in a manner that cannot be captured clearly by evaluating ensemble-averaged readouts like secondary structure propensities. Cut-based FEPs represent an appropriate way to condense trajectory information and to show approximately the distributions of barriers and basins. In Fig. 3, free Aβ_{12-28} exhibited no predominantly populated conformation. When clustering data using the Cα atoms of residues 14-24, the most populated conformer corresponds to a straight helical structure (see Fig. 7) with a disordered C-terminus. The entire helical basin encompasses partially helical as well as completely helical structures that constitute a total statistical weight of about 10%. It is separated by a wide barrier region comprised of minor basins with partially helical states (population of about 10%) from a broad entropic basin (70%). The latter consists of fluctuating conformers devoid of regular secondary structure content. Figs. S2-S11 in Supplemental Data show the same type of data for simulations in the presence of inhibitors. In all cases, a large entropic basin has the largest statistical weight indicating that inhibitors are not able to lock Aβ_{12-28} into a specific conformation. There are changes, however, to the relative weights of ordered conformers. In particular, a basin is increasingly populated that corresponds to a compact structure characterized by a specific loop conformation spanning residues 14-24. It is stabilized (see Fig. 7) by the formation of a hydrogen-bonded network of side chains including Asp23, His13 and either Gln15 or His14. Its statistical weight for each inhibitor is summarized in Table S1 in Supplemental Data and ranges from ca. 2 to 20%. Given that the population of the loop is lowest for the free system and in the presence of largely inert β-Ala-His, and that it is only transiently populated in either case (see Fig. S12-S13 in Supplemental Data), one may ask how a specific Aβ_{12-28} conformation can be universally enhanced by other inhibitors in the absence of a specific binding mode.

The answer is that the inhibitors provide an

additional non-aqueous interface, *i.e.*, they allow patterns of sequestering residues from solvent to change. Normalized by their propensity to bind Aβ_{12-28}, the indole-containing inhibitors NQTrp (see Table S1 in Supplemental Data) and ^{D}Trp-Aib exert the largest effect in this regard. Indole groups have been heavily implicated in Aβ aggregation inhibitor design (39), and were also analyzed in systematic fashion as inhibitors of hen egg white lysozyme aggregation (41). Interestingly, the inhibitor is not required to interact directly with the polar residues mentioned, but can act indirectly or allosterically. Such a mechanism of action is entropically favored on account of the peptide's high intrinsic flexibility (this is true for example for ^{D}Trp-Aib as seen in see Fig. S14 in Supplemental Data). However, Fig. S14 also shows a counterexample: Direct binding of anthracene to the side chains of Asp23 and Gln15 occurs with greater frequency than binding in general, and population of the loop conformer in the absence of binding is essentially identical to that seen for the free system. These data resemble a conformational selection mechanism described in binding equilibria of intrinsically disordered proteins (82).

An interfacial effect as described is intriguing as it allows for an explanation as to why so many different inhibitors, which share some structural hallmarks have been characterized (83,84). If one compares the requirements for small molecule design in enzyme inhibition to that of inhibition of oligomerization and fibrillization, one would be forced to conclude that the intrinsic disorder of the target translates to the inhibitor level. This could imply that "poly-functional" moieties combining largely hydrophobic parts with polar sites are well-suited to bind disordered Aβ. Disordered interactions of such poly-functional units with peptide moieties may also explain why mutation studies have failed to establish the necessity for aromatic residues Phe19 and Phe20 in Aβ to be present for aromatic inhibitors of the general flavor tested here to be effective (85).

It should be noted that the partial desolvation of a charged moiety (Asp23) may raise questions toward the accuracy of the computational model in use. We wish to emphasize that we assign little importance to the structural details of the aforementioned loop structure on account of these concerns, even though desolvation of Asp23 is seen experimentally in the NMR structures (86) and in explicit solvent simulations (87) of similar fragments. It is worthwhile to point out that in our simulations the statistical weight of the loop never exceeds 20% meaning that most computed ensemble averages would only be moderately affected.

Comparison with NMR NOE data. Aside from keeping the project computationally tractable, an added motivation for studying the specific fragment Aβ_{12-28} was the availability of NMR spectroscopic data in the presence of different inhibitors. For the inhibitor Tyr-Aib-Trp-Phe, Frydman-Marom *et al.* report a few intermolecular NOEs indicating preferential binding of Aib to residues 20-22 (76). Undoubtedly, this is qualitatively compatible with our data as presented in Fig. 4. For Aβ_{12-28} in the presence of NQTrp (4:1 excess of Aβ_{12-28}), Scherzer-Attali *et al.* (43) report 8 long-range NOEs. Four of them cover a sequence spacing of only *i,i+4*, and five of them involve Val18. However, with the exception of an NOE between the α-proton of Asp23 and the backbone amide proton of Gln15, all of them are weak. When computing predicted NOEs from simulation data, we found a significant number of long-range NOEs that should have been detectable by NMR irrespective of whether we use 3rd or 6th power averaging (88). This presence of "false positives" indicates either incompatible ensembles, NMR-intrinsic issues when studying disordered systems, *e.g.*, issues pertaining to the lifetime of conformational states (89), or inapplicability of the model used for prediction, considering that most hydrogen positions had to be modeled *a posteriori*. Coupled to the fact that the source of NOE signals at a 4:1 excess concentration of Aβ_{12-28} can hardly be expected to delineate Aβ-intrinsic and inhibitor-induced signals, we conclude that a quantitative comparison between measured NMR parameters and simulation data carries more caveats than potential. This is particularly true given the lack of reliable long-range NOEs observed for Aβ_{12-28} alone (77).

Structural ensembles derived from NMR data suggest that Aβ_{12-28} adopts loop-like or hairpin-like structures devoid of canonical backbone hydrogen bonds (40,43), and this trend is indirectly supported by independent studies on different fragments (86,90,91). Interestingly, the

aforementioned loop state that is stabilized in the presence of inhibitors features a prominent contact between residues Gln15 and Asp23 (see Fig. 4). In our simulations, it is however the amide hydrogens of the glutamine side chain that mediate this contact (see Fig. 7B). Lastly, due to the very weak impact of inhibitors on average secondary structure populations (see Fig. 6), we do not attempt to correlate those changes to chemical shift differences observed in the presence of inhibitors.

CONCLUSIONS

The FACTS implicit solvation model allowed us to obtain data with statistical errors that permit establishing subtle, quantitative effects of inhibitor binding (for example, Fig. 6). Therefore, five main results emerge from the comparative analysis of MD simulations of Aβ_{12-28} alone and in the presence of small molecules that have been shown experimentally to interfere with Aβ peptide self-assembly.

1. The free energy surface of Aβ_{12-28} shows a broad entropic basin devoid of a predominant conformation (Fig. 3 and Figs. S1-S11 in Supplemental Data). Upon inhibitor binding, significant, but subtle changes to the free energy landscape occur as exemplified by changes in average secondary structure propensities and intramolecular contacts (Figs. 4 and 6).
2. Different inhibitors share common intermolecular interactions but none of them exhibits a specific binding mode irrespective of binding affinity (Table 1 and Fig. 4).
3. Differences in affinity can be partially rationalized by the relative number of aromatic groups and charged groups in the inhibitors (Table 1). The former provide a favorable contribution through hydrophobic interactions mainly with Phe19 and Phe20 and through mixed polar and nonpolar interactions with residues 13-18. The desolvation of charged groups is generally unfavorable, *i.e.*, inhibitor binding does not rely on salt bridge formation (Figs. 4 and 5).
4. The simulation data demonstrate that detailed predictions of interactions at the molecular level cannot just rely on simple heuristics (92). As an example, the relative binding properties of the different endomorphin variants appear nearly impossible to rationalize with purely sequence-based approaches (Table 1).
5. The data show how small molecules composed of similar building blocks can incur a generic interfacial effect. Such effect is the cause of the enhancement of a specific loop conformation of Aβ_{12-28} (Table S1 and Fig. S13 in Supplemental Data).

Taken together, the results show how aggregation inhibitors can have subtle, but significant effects on the behavior of a fragment of Aβ encompassing the CHC that is most commonly implied in the peptide's amyloidogenicity. It is important to ask whether our results could apply to other aggregation-prone IDPs as well. We use α-synuclein, the disease protein of Parkinson's, as an example, which is significantly larger than Aβ, and adopts more extended conformations in solution (93). Despite large differences in sequence properties, monomeric ensembles, and oligomer distributions, there exists a comparable range of compounds possessing one or more aromatic moieties that have been shown to bind monomeric α-synuclein (94), and to interfere with α-synuclein aggregation (53,95-98) and oligomerization (99,100). It is not necessary and perhaps unlikely that all these inhibitors act in equivalent fashion for diverse systems. Returning to Fig. 1, this means for example that the multiple equilibria involving inhibitor binding may not all contribute significantly. As an example, our data for β-Ala-His suggest that this compound is unlikely to interfere with Aβ dimerization at the low concentrations typically in use in vitro. Similarly, ^{D}Trp-Aib exhibited no significant effects on early oligomer formation of Aβ_{1-42} at concentration ratios of up to 40:1 excess of ^{D}Trp-Aib (40), and is also largely unbound in our simulations. This suggests a specific role in step VI for those two compounds. Conversely, evidence for generic effects at the oligomer level (in particular step V in Fig. 1) is not only given in the wider context of enzyme activity assays (101) but has also been proposed specifically in the context of aggregation experiments (102).

Finally, the consistency between simulation results and experimental data (NMR spectroscopy analysis of Aβ_{12-28} and characterization of inhibition of Aβ_{40} aggregation) suggest that several of the conclusions are independent of the details of

the simulation model. Encouraged by these results, we are currently using the same methodology to perform a medium-throughput screening of a novel class of compounds meant to interfere with Aβ oligomerization and fibrillization.

REFERENCES

1. Bharadwaj, P. R., Dubey, A. K., Masters, C. L., Martins, R. N., and Macreadie, I. G. (2009) *J. Cell. Mol. Med.* **13**, 412-421
2. Selkoe, D. J. (1991) *Neuron* **6**, 487-498
3. Terry, R. D. (1994) *Prog. Brain Res.* **101**, 383-390
4. Lomakin, A., Chung, D. S., Benedek, G. B., Kirschner, D. A., and Teplow, D. B. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 1125-1129
5. Fawzi, N. L., Okabe, Y., Yap, E.-H., and Head-Gordon, T. (2007) *J. Mol. Biol.* **365**, 535-550
6. Pellarin, R., and Caflisch, A. (2006) *J. Mol. Biol.* **360**, 882-892
7. Rochet, J. C., and Lansbury, P. T. (2000) *Curr. Opin. Struct. Biol.* **10**, 60-68
8. Serio, T. R., Cashikar, A. G., Kowal, A. S., Sawicki, G. J., Moslehi, J. J., Serpell, L., Arnsdorf, M. F., and Lindquist, S. L. (2000) *Science* **289**, 1317-1321
9. Wu, C., and Shea, J. E. (2011) *Curr. Opin. Struct. Biol.* **21**, 209-220
10. Kim, S., Takeda, T., and Klimov, D. K. (2010) *Biophys. J.* **99**, 1949-1958
11. Tarus, B., Straub, J. E., and Thirumalai, D. (2005) *J. Mol. Biol.* **345**, 1141-1156
12. Thirumalai, D., Tarus, B., and Straub, J. E. (2008) *J. Mol. Biol.* **379**, 815-829
13. Lansbury, P. T., and Lashuel, H. A. (2006) *Nature* **443**, 774-779
14. Haass, C., and Selkoe, D. J. (2007) *Nat. Rev. Mol. Cell Bio.* **8**, 101-112
15. Uversky, V. N., Oldfield, C. J., Midic, U., Xie, H., Xue, B., Vucetic, S., Iakoucheva, L. M., Obradovic, Z., and Dunker, A. K. (2009) *BMC Genomics* **10**, S7
16. Uversky, V. N., Oldfield, C. J., and Dunker, A. K. (2008) *Annu. Rev. Biophys.* **37**, 215-246
17. Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., and Jones, D. T. (2004) *J. Mol. Biol.* **337**, 635-645
18. Mittag, T., and Forman-Kay, J. D. (2007) *Curr. Opin. Struct. Biol.* **17**, 3-14
19. Uversky, V., Gillespie, J. R., and Fink, A. L. (2000) *Proteins: Struct. Func. Genet.* **41**, 415-427
20. Dunker, A. K., and Obradovic, Z. (2001) *Nat. Biotechnol.* **19**, 805-806
21. Vitalis, A., and Caflisch, A. (2010) *J. Mol. Biol.* **403**, 148-165
22. Chen, Y., and Glabe, C. G. (2006) *J. Biol. Chem.* **281**, 24414-24422
23. Wright, P. E., and Dyson, H. J. (2009) *Curr. Opin. Struct. Biol.* **19**, 31-38
24. Pappu, R. V., Wang, X., Vitalis, A., and Crick, S. L. (2008) *Arch. Biochem. Biophys.* **469**, 132-141
25. Linding, R., Schymkowitz, J., Rousseau, F., Diella, F., and Serrano, L. (2004) *J. Mol. Biol.* **342**, 345-353
26. Fields, G. B., Alonso, D. O. W., Stigter, D., and Dill, K. A. (1992) *J. Phys. Chem.* **96**, 3974-3981
27. Cohen, F. E., and Kelly, J. W. (2003) *Nature* **426**, 905-909
28. Necula, M., Kayed, R., Milton, S., and Glabe, C. G. (2007) *J. Biol. Chem.* **282**, 10311-10324
29. Kokkoni, N., Stott, K., Amijee, H., Mason, J. M., and Doig, A. J. (2006) *Biochemistry* **45**, 9906-9918
30. Yan, L. M., Velkova, A., Tatarek-Nossol, M., Andreetto, E., and Kapurniotu, A. (2007) *Angew. Chem. Int. Ed.* **46**, 1246-1252
31. Ban, T., Hoshino, M., Takahashi, S., Hamada, D., Hasegawa, K., Naiki, H., and Goto, Y. (2004) *J. Mol. Biol.* **344**, 757-767
32. Kanapathipillai, M., Lentzen, G., Sierks, M., and Park, C. B. (2005) *FEBS Lett.* **579**, 4775-4780
33. Gervais, F., Paquette, J., Morissette, C., Krzywkowski, P., Yu, M., Azzi, M., Lacombe, D., Kong, X. Q., Aman, A., Laurin, J., Szarek, W. A., and Tremblay, P. (2007) *Neurobiol. Aging* **28**, 537-547
34. Mishra, R., Bulic, B., Sellin, D., Jha, S., Waldmann, H., and Winter, R. (2008) *Angew. Chem. Int. Ed.* **47**, 4679-4682
35. Porat, Y., Mazor, Y., Efrat, S., and Gazit, E. (2004) *Biochemistry* **43**, 14454-14462

36. Bastianetto, S., Krantic, S., and Quirion, R. (2008) *Mini Rev. Med. Chem.* **8**, 429-435
37. Convertino, M., Pellarin, R., Catto, M., Carotti, A., and Caflisch, A. (2009) *Protein Sci.* **18**, 792-800
38. Derreumaux, P., and Chebaro, Y. (2009) *Proteins: Struct. Funct. Bioinform.* **75**, 442-452
39. Cohen, T., Frydman-Marom, A., Rechter, M., and Gazit, E. (2006) *Biochemistry* **45**, 4727-4735
40. Frydman-Marom, A., Rechter, M., Shefler, I., Bram, Y., Shalev, D. E., and Gazit, E. (2009) *Angew. Chem. Int. Ed.* **48**, 1981-1986
41. Morshedi, D., Rezaei-Ghaleh, N., Ebrahim-Habibi, A., Ahmadian, S., and Nemat-Gorgani, M. (2007) *FEBS J* **274**, 6415-6425
42. Pickhardt, M., Gazova, Z., von Bergen, M., Khlistunova, I., Wang, Y., Hascher, A., Mandelkow, E. M., Biernat, J., and Mandelkow, E. (2005) *J. Biol. Chem.* **280**, 3628-3635
43. Scherzer-Attali, R., Pellarin, R., Convertino, M., Frydman-Marom, A., Egoz-Matia, N., Peled, S., Levy-Sakin, M., Shalev, D. E., Caflisch, A., Gazit, E., and Segal, D. (2010) *PLoS One* **5**, e11101
44. Jiang, P., Li, W. F., Shea, J. E., and Mu, Y. G. (2011) *Biophys. J.* **100**, 2076-2076
45. Evers, F., Jeworrek, C., Tiemeyer, S., Weise, K., Sellin, D., Paulus, M., Struth, B., Tolan, M., and Winter, R. (2009) *J Am Chem Soc* **131**, 9516-9521
46. Bieschke, J., Russ, J., Friedrich, R. P., Ehrnhoefer, D. E., Wobst, H., Neugebauer, K., and Wanker, E. E. (2010) *Proc. Natl. Acad. Sci. U. S. A.* **107**, 7710-7715
47. Takeda, T., Kumar, R., Raman, E. P., and Klimov, D. K. (2010) *J. Phys. Chem. B* **114**, 15394-15402
48. Kim, S., Chang, W. E., Kumar, R., and Klimov, D. K. (2011) *Biophys. J.* **100**, 2024-2032
49. Landau, M., Sawaya, M. R., Faull, K. F., Laganowsky, A., Jiang, L., Sievers, S. A., Liu, J., Barrio, J. R., and Eisenberg, D. (2011) *PLoS Biol.* **9**, e1001080
50. Sievers, S. A., Karanicolas, J., Chang, H. W., Zhao, A., Jiang, L., Zirafi, O., Stevens, J. T., Munch, J., Baker, D., and Eisenberg, D. (2011) *Nature* **475**, 96-100
51. Schutz, A. K., Soragni, A., Hornemann, S., Aguzzi, A., Ernst, M., Bockmann, A., and Meier, B. H. (2011) *Angew. Chem. Int. Ed.* **50**, 5956-5960
52. Lorenzo, A., and Yankner, B. A. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 12243-12247
53. Braga, C. A., Follmer, C., Palhano, F., Khattar, E., Freitas, M. S., Romão, L., Di Giovanni, S., Lashuel, H. A., Silva, J. L., and Foguel, D. (2011) *J. Mol. Biol.* **405**, 254-273
54. De Felice, F., Vieira, M. N. N., Saraiva, L. M., Figueroa-VIllar, J. D., Garcia-Abreu, J., Liu, R., Chang, L., Klein, W. L., and Ferreira, S. T. (2004) *FASEB J.* **18**, 1366-1372
55. Riek, R., Güntert, P., Döbeli, H., Wipf, B., and Wüthrich, K. (2001) *Eur. J. Biochem.* **268**, 5930-5936
56. Cohlberg, J. A., Li, J., Uversky, V. N., and Fink, A. L. (2002) *Biochemistry* **41**, 1502-1511
57. Feng, Y., Yang, S.-G., Du, X.-T., Zhang, X., Sun, X.-X., Zhao, M., Sun, G.-Y., and Liu, R.-T. (2009) *Biochem. Biophys. Res. Commun.* **390**, 1250-1254
58. Straub, J. E., Guevara, J., Huo, S., and Lee, J. P. (2002) *Acc. Chem. Res.* **35**, 473-481
59. Zhang, S., Casey, N., and Lee, J. P. (1998) *Fold Des.* **3**, 413-422
60. Porat, Y., Abramowitz, A., and Gazit, E. (2006) *Chem. Biol. Drug Des.* **67**, 27-37
61. Tjernberg, L. O., Naslund, J., Lindqvist, F., Johansson, J., Karlstrom, A. R., Thyberg, J., Terenius, L., and Nordstedt, C. (1996) *J. Biol. Chem.* **271**, 8545-8548
62. Williams, A. D., Portelius, E., Kheterpal, I., Guo, J. T., Cook, K. D., Xu, Y., and Wetzel, R. (2004) *J. Mol. Biol.* **335**, 833-842
63. Krivov, S. V., and Karplus, M. (2006) *J. Phys. Chem. B* **110**, 12689-12698
64. Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009) *J. Comput. Chem.* **30**, 1545-1614
65. Haberthür, U., and Caflisch, A. (2008) *J. Comput. Chem.* **29**, 701-715
66. Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) *J Chem. Theory Comput.* **4**, 435-447
67. Bussi, G., Donadio, D., and Parrinello, M. (2007) *J. Chem. Phys.* **126**
68. Parrinello, M., and Rahman, A. (1981) *J. Appl. Phys.* **52**, 7182-7190
69. Buck, M., Bouguet-Bonnet, S., Pastor, R. W., and MacKerell, A. D. (2006) *Biophys. J.* **90**, L36-L38

70. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) *J. Chem. Phys.* **79**, 926-935
71. Darden, T., York, D., and Pedersen, L. (1993) *J. Chem. Phys.* **98**, 10089-10092
72. Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) *J. Comput. Chem.* **18**, 1463-1472
73. Seeber, M., Felline, A., Raimondi, F., Muff, S., Friedman, R., Rao, F., Caflisch, A., and Fanelli, F. (2011) *J. Comput. Chem.* **32**, 1183-1194
74. Vitalis, A., Wang, X. L., and Pappu, R. V. (2007) *Biophys. J.* **93**, 1923-1937
75. Schäfer, L. (1999) *Excluded Volume Effects in Polymer Solutions: as Explained by the Renormalization Group.*, 1st ed., Springer, Berlin, Germany
76. Frydman-Marom, A., Convertino, M., Pellarin, R., Lampel, A., Shaltiel-Karyo, R., Segal, D., Caflisch, A., Shalev, D. E., and Gazit, E. (September 19, 2011) *ACS Chem. Biol.*, 10.1021/cb200103h
77. Attanasio, F., Convertino, M., Caflisch, A., Corazza, A., Esposito, G., Cataldo, S., Pignataro, B., Milardi, D., and Rizzarelli, E. (2011) *J. Phys. Chem. B* **in review**
78. Tartaglia, G. G., Cavalli, A., Pellarin, R., and Caflisch, A. (2005) *Protein Sci.* **14**, 2723-2734
79. Caflisch, A. (2006) *Curr. Opin. Chem. Biol.* **10**, 437-444
80. Wurth, C., Guimard, N. K., and Hecht, M. H. (2002) *J. Mol. Biol.* **319**, 1279-1290
81. Azriel, R., and Gazit, E. (2001) *J. Biol. Chem.* **276**, 34156-34161
82. Csermely, P., Palotai, R., and Nussinov, R. (2010) *Trends Biochem. Sci* **35**, 539-546
83. Hawkes, C. A., Ng, V., and McLaurin, J. (2009) *Drug Dev. Res.* **70**, 111-124
84. Re, F., Airoldi, C., Zona, C., Masserini, M., La Ferla, B., Quattrocchi, N., and Nicotra, F. (2010) *Curr. Med. Chem.* **17**, 2990-3006
85. Armstrong, A. H., Chen, J., McKoy, A. F., and Hecht, M. H. (2011) *Biochemistry* **50**, 4058-4067
86. Zhang, S., Iwata, K., Lachenmann, M. J., Peng, J. W., Li, S., Stimson, E. R., Lu, Y., Felix, A. M., Maggio, J. E., and Lee, J. P. (2000) *J. Struct. Biol.* **130**, 130-141
87. Baumketner, A., Krone, M. G., and Shea, J. E. (2008) *Proc. Natl. Acad. Sci. U. S. A.* **105**, 6027-6032
88. Zagrovic, B., and van Gunsteren, W. F. (2006) *Proteins: Struct. Funct. Bioinform.* **63**, 210-218
89. Eliezer, D. (2009) *Curr. Opin. Struct. Biol.* **19**, 23-30
90. Lazo, N. D., Grant, M. A., Condron, M. C., Rigby, A. C., and Teplow, D. B. (2005) *Protein Sci.* **14**, 1581-1596
91. Sgourakis, N. G., Yan, Y. L., McCallum, S. A., Wang, C. Y., and Garcia, A. E. (2007) *J. Mol. Biol.* **368**, 1448-1457
92. Stempler, S., Levy-Sakin, M., Frydman-Marom, A., Amir, Y., Scherzer-Attali, R., Buzhansky, L., Gazit, E., and Senderowitz, H. (2011) *J. Comput. Aided Mol. Des.* **25**, 135-144
93. Uversky, V. N. (2007) *J. Neurochem.* **103**, 17-37
94. Latawiec, D., Herrera, F., Bek, A., Losasso, V., Candotti, M., Benetti, F., Carlino, E., Kranjc, A., Lazzarino, M., Gustincich, S., Carloni, P., and Legname, G. (2010) *PLoS One* **5**, e9234
95. Dong-Pyo, H., Uversky, V. N., and Fink, A. L. (2008) *J. Mol. Biol.* **383**, 214-223
96. Zhou, W., Gallagher, A., Dong-Pyo, H., Long, C., Fink, A. L., and Uversky, V. N. (2009) *J. Mol. Biol.* **388**, 597-610
97. Kim, J., Harada, R., Kobayashi, M., Kobayashi, N., and Sode, K. (2010) *Mol. Neurodegener.* **5**, 20
98. Pandey, N., Strider, J., Nolan, W. C., Yan, S. X., and Galvin, J. E. (2008) *Acta Neuropathol.* **115**, 479-489
99. Caruana, M., Högen, T., Levin, J., Hillmer, A. S., Giese, A., and Vassallo, N. (2011) *FEBS Lett.* **585**, 1113-1120
100. Hillmer, A. S., Putcha, P., Levin, J., Högen, T., Hyman, B. T., Kretzschmar, H., McLean, P. J., and Giese, A. (2010) *Biochem. Biophys. Res. Commun.* **391**, 461-466
101. Shoichet, B. K. (2006) *Drug Discov. Today* **11**, 607-615
102. Feng, B. Y., Toyama, B. H., Wille, H., Colby, D. W., Collins, S. R., May, B. C. H., Prusiner, S. B., Weissman, J., Shoichet, B. K. (2008) *Nat. Chem. Biol.* **4**, 197-199
103. Pappu, R. V., and Tran, H. T. (2006) *Biophys. J.* **91**, 1868-1886

104. Carter, P., Andersen, C. A. F., and Rost, B. (2003) *Nucleic Acids Res.* **31**, 3293-3295
105. Szegedi, V., Juhasz, G., Rozsa, E., Juhasz-Vedres, G., Datki, Z., Fulop, L., Bozso, Z., Lakatos, A., Laczko, I., Farkas, T., Kis, Z., Toth, G., Soos, K., Zarandi, M., Denes, B., Toldi, J., and Penke, B. (2006) *FASEB J.* **20**, 1191-1193

Acknowledgments-We thank Dr. Riccardo Pellarin, Andrea Magno, and François Marchand for interesting discussions and suggestions. Most of the simulations were carried out on the Schrödinger cluster at the Informatikdienste of the University of Zurich.

FIGURE LEGENDS

Fig. 1. **Schematic depicting the coupled equilibria that could be involved in inhibition of fibrillization and/or oligomerization of Aβ by small molecules**. Largely unstructured peptides are shown as gray ribbons with the CHC highlighted in yellow. Peptides in fibril-compatible conformations are shown in red, and inhibitor molecules in blue. Angular brackets indicate conformational equilibria at different assembly levels, and curly brackets a repeating unit replicated along the indicated axis. Steps I and II pertain to the unperturbed, nucleation-dependent aggregation pathway. On- and off-pathway assembly steps beyond the dimer are not shown explicitly. Steps VII and VIII are the analogous steps with inhibitor bound to the aggregating peptides. Finally, steps III-VI describe binding equilibria to various peptide species. Preferential interactions of the CHC with inhibitor molecules containing aromatic moieties have been postulated (81). See text for further details.

Fig. 2. **Average polymeric properties for Aβ12-28 alone and in the presence of various inhibitors.** Indicated errors are minimum and maximum values obtained from partitioning the data into six blocks. Panel A shows the average atom-to-atom distance for pairs of residues as a function of sequence separation. Capping groups are included as separate residues in this analysis. Collapsed globules would be indicated by the internal distance reaching a plateau value for sequence spacings beyond the length scale of local rigidity. Random coil-like (good solvent) chains are significantly more extended as indicated by the black, dotted line [data from simulation of Aβ12-28 in the excluded volume limit (103)]. Panel B shows the angular correlation function of N→C vectors as a function of sequence separation. Negative values indicate the chain turning on itself, *i.e.*, (partial) collapse. A good solvent chain leads to simple, monotonous decay of the correlation function (black dotted line).

Fig. 3. **Monomeric Aβ12-28 is mainly unstructured.** (Top) DSSP analysis of the sampling arranged according to the reaction coordinate of the cFEP (the color-code is: helix in red; β-extended in green; loop and turn in blue; bend is the DSSP letter S and is plotted in yellow for clarity). (Bottom) The cFEP of monomeric Aβ12-28 using its most populated conformer as the reference state (the value of the reaction coordinate is zero for the reference state). Representative conformations of individual free-energy basins are reported as insets (residues 14-24 are highlighted in yellow in the cartoons).

Fig. 4. **Intra- and intermolecular contact maps of Aβ12-28 alone and in the presence of various inhibitors.** The color legend applies to all panels. Axes are labeled with single-letter amino acids codes. "B" stands for β-alanine, "DT" for ^{D}Trp, "NQ" for naphtoquinone, "Z" for α-aminoisobutyric acid (Aib), "AC" for anthracene, and "AQ" for anthraquinone. Capping groups for peptides are considered as separate residues, but are not labeled on the axes. A contact is defined as any two atoms of the corresponding residues being separated by less than 5.5 Å. The diagonal is excluded. Upper left half matrices contain average contact probabilities, and lower right half matrices the corresponding standard errors estimated as half the difference between the minimum and maximum values measured over six blocks. The region denoted by blue lines shows a specific contact pattern corresponding to the loop structure discussed in the text.

Fig. 5. **Matrix of van der Waals interaction energy between NQTrp and Aβ12-28**. The interaction energy between individual functional groups of NQTrp and Aβ12-28 residues (backbone and side chain atoms) was computed by CHARMM using every 100th snapshot. NQTrp was decomposed into single functional groups with net integer charges (abbreviations CO and CO´, COO^{-}, NH, indole, and NQ stand for quinonic carbonyls, carboxy, amide, indole and naphtoquinone moieties, respectively [see Table 1 for the chemical structure of NQTrp]). The sum of all pairwise averages reported in the matrix is -13.3 kcal/mol.

Fig. 6. **Influence of different inhibitors of aggregation on secondary structure propensities of monomeric Aβ12-28.** The secondary structure content was calculated by the DSSPcont algorithm (104) as implemented in Wordom (73). (Top) "Loop and turn" are DSSP letters L, S, and T. (Middle) "β-extended" consists of DSSP letters E and B. (Bottom) "Helical" are DSSP letters G, H, and I. The error bars indicate the range between the maximum and minimum value over three independent 5 μs simulations.

Fig. 7. **Helical and loop conformations identified with cFEP.** The helical conformation is shown in panel (A). The loop conformation is stabilized by the hydrogen bonds between the side chain of Asp23 and backbone NH groups as well as the side chains of either Gln15 (B) or His14 (C).

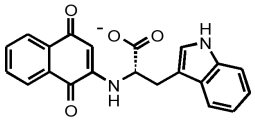
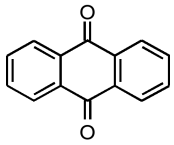
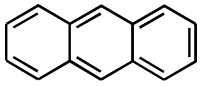
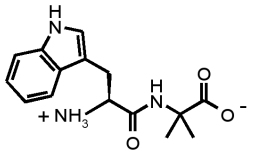
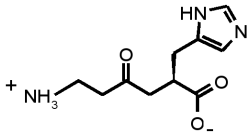
Inhibitor [a]	F_b to Aβ12-28 [b]	Effective Binding Energy				Ref.	Structure [e]
		vdW [c]	El [c]	Np [c]	Total [d]		
NQTrp	71	-15.8	4.8	-3.0	-13.9	(43)	
9,10-anthraquinone	39	-6.6	2.5	-1.5	-5.6	(37)	
anthracene	67	-7.7	1.9	-1.7	-7.5	(37)	
Tyr-Aib-Trp-Phe	51	-9.6	3.2	-2.8	-9.2	(76)	
Tyr-Pro-Trp-Phe	50	-10.3	3.4	-2.9	-9.8	(76)	
Tyr-Pro-Phe-Phe	55	-16.3	4.3	-3.6	-15.6	(76)	
Tyr-Pro-Trp-Phe-NH_2	62	-11.9	4.1	-3.3	-11.1	/	
Tyr-Pro-Phe-Phe-NH_2	50	-10.5	3.5	-3.0	-10.0	(105)	
[D]Trp-Aib	19	-4.5	1.8	-1.8	-4.5	(40)	
β-Ala-His	10	-0.9	0.9	-0.8	-0.8	(77)	

Table 1: Small-molecule inhibitors and their affinity to Aβ12-28

[a]Aib: α-aminoisobutyric acid; NQTrp: 1,4-napthoquinon-2-yl-L-tryptophan; Tyr-Aib-Trp-Phe: designed derivative of endomorphin-1; Tyr-Pro-Trp-Phe: endomorphin-1 with charged carboxy group at C-terminus; Tyr-Pro-Phe-Phe: endomorphin-2 with charged carboxy group at C-terminus; Tyr-Pro-Trp-Phe-NH_2: endomorphin-1; Tyr-Pro-Phe-Phe-NH_2: endomorphin-2; β-Ala-His: carnosine. Note that the endogenous endomorphins have the $–NH_2$ group at the C-terminus.

[b]F_b is the fraction of snapshots in which any atom of the inhibitor is within 7.5 Å of any atom of Aβ12-28. Values are given in percent.

[c]Each of the these three terms (vdW = van der Waals terms; El = electrostatic interactions; Np = nonpolar solvation terms) was calculated by subtracting the energy of Aβ12-28 and inhibitor in the isolated state from the energy in the ensemble of bound conformations.

[d]The effective binding energy is the sum of the changes in the van der Waals energy, electrostatic energy, and non-polar solvation energy upon binding. All values are reported in kcal/mol. They include all contributions to the binding free energy except for changes in entropies of the solutes.

[e]Structure diagrams are omitted for tetrapeptides.

Figure 1

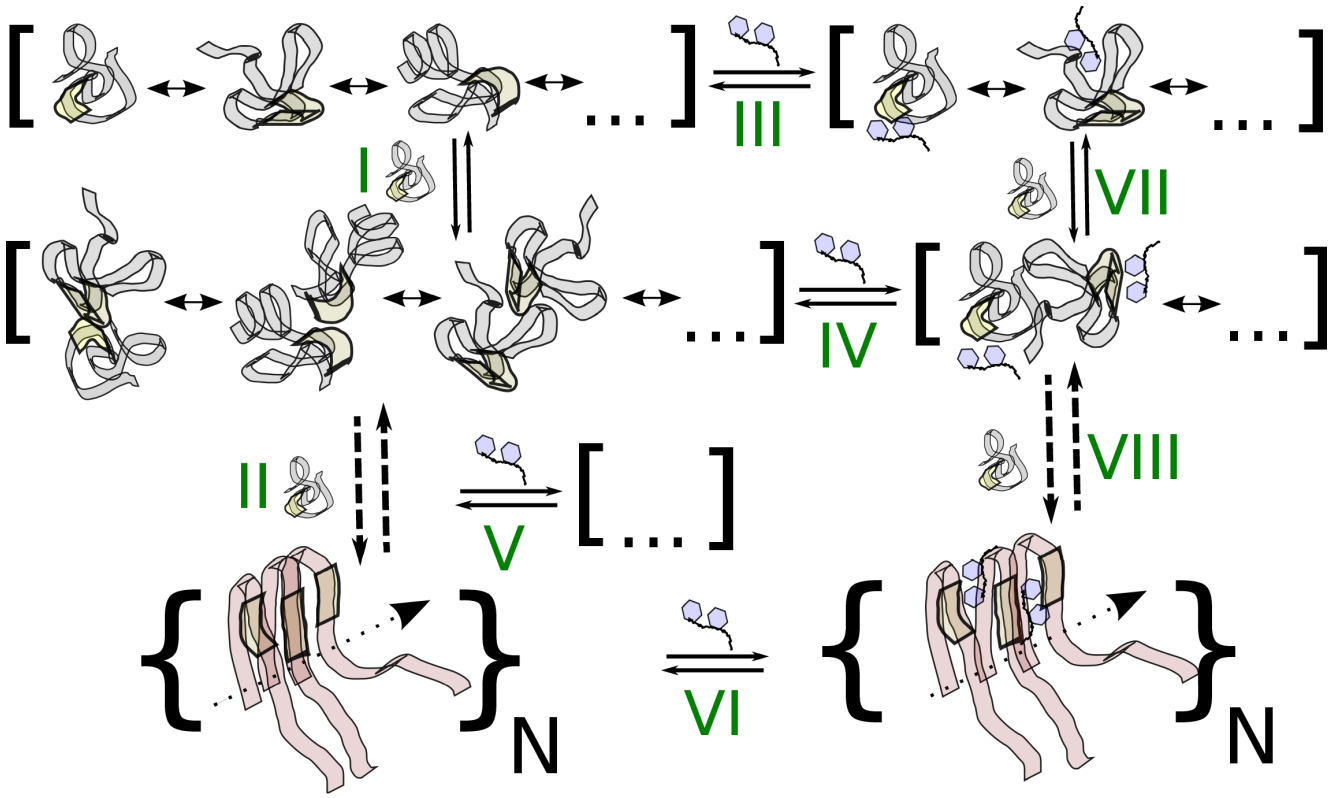


Figure 2

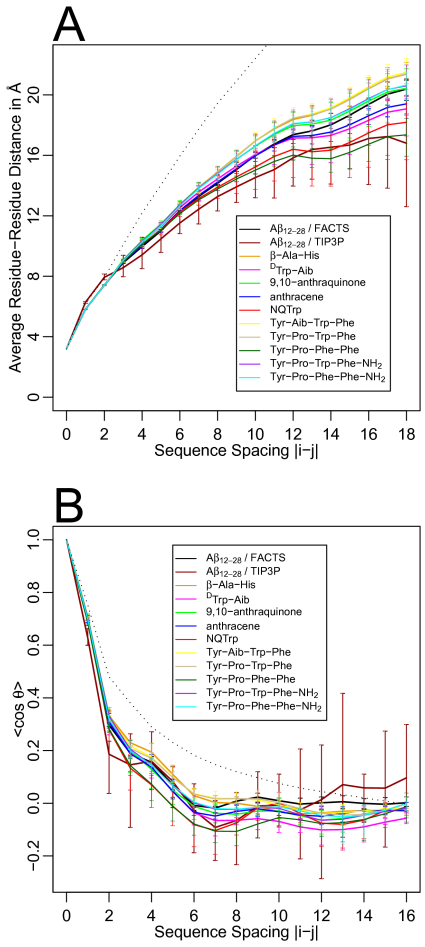


Figure 3

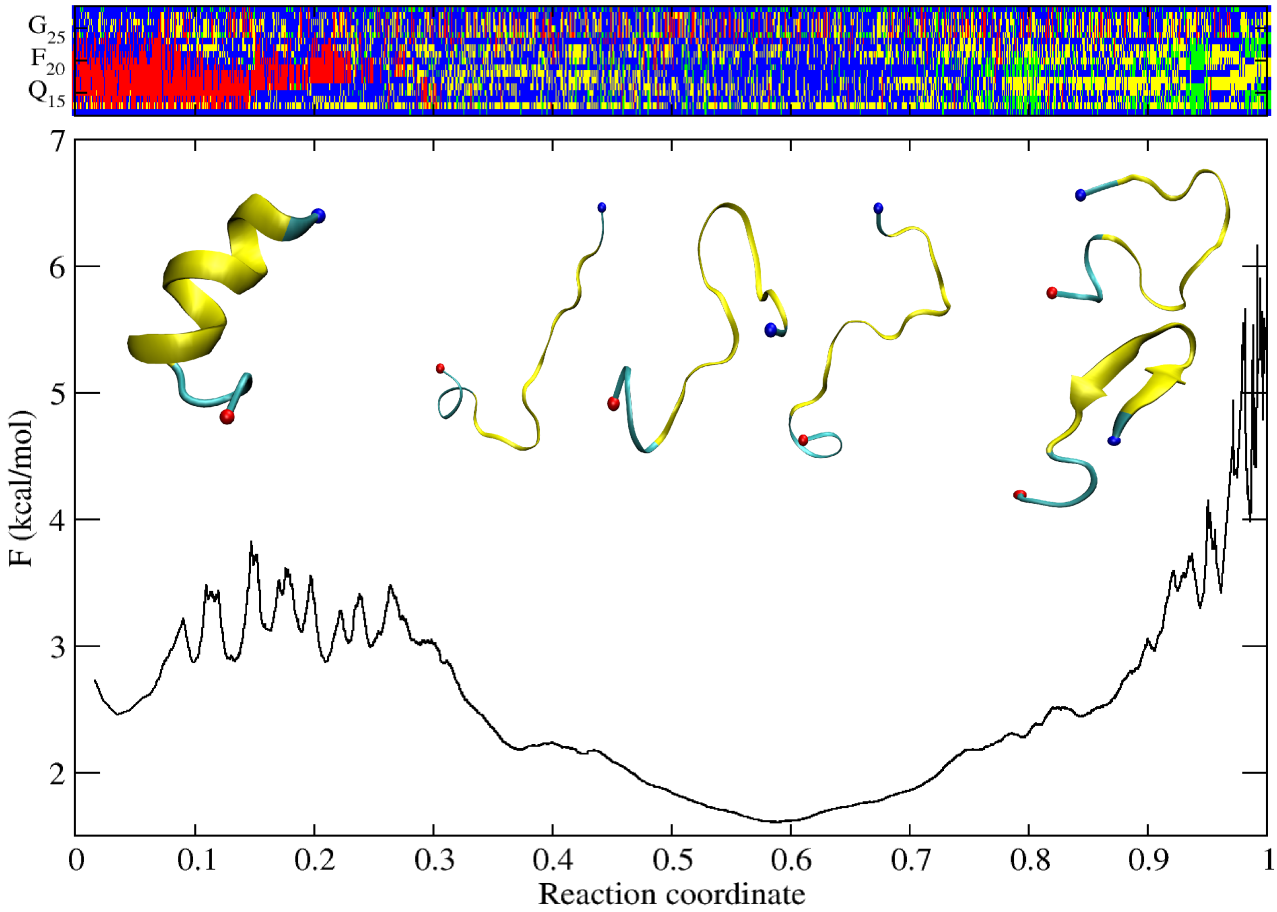


Figure 4

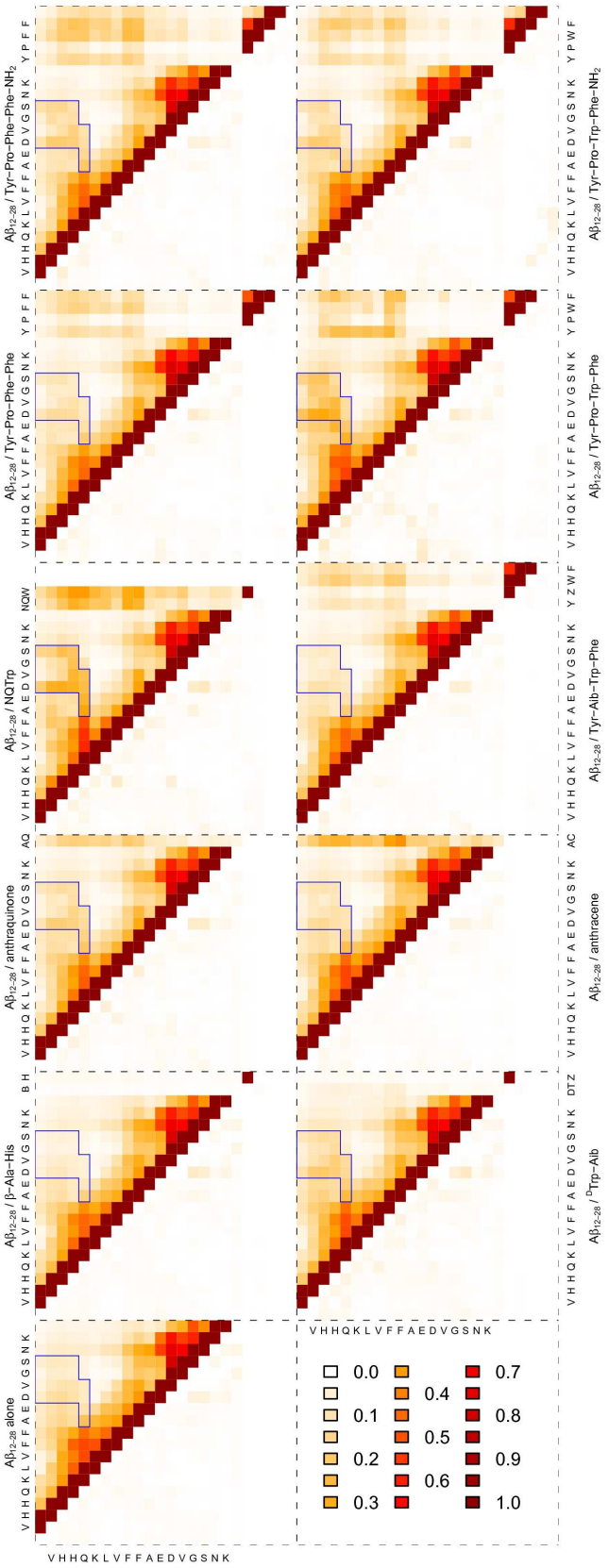


Figure 5

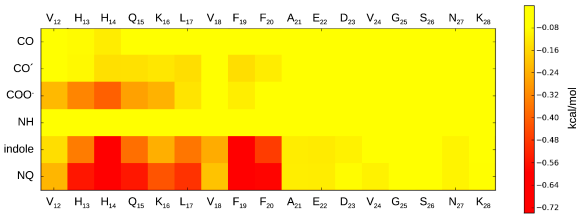


Figure 6

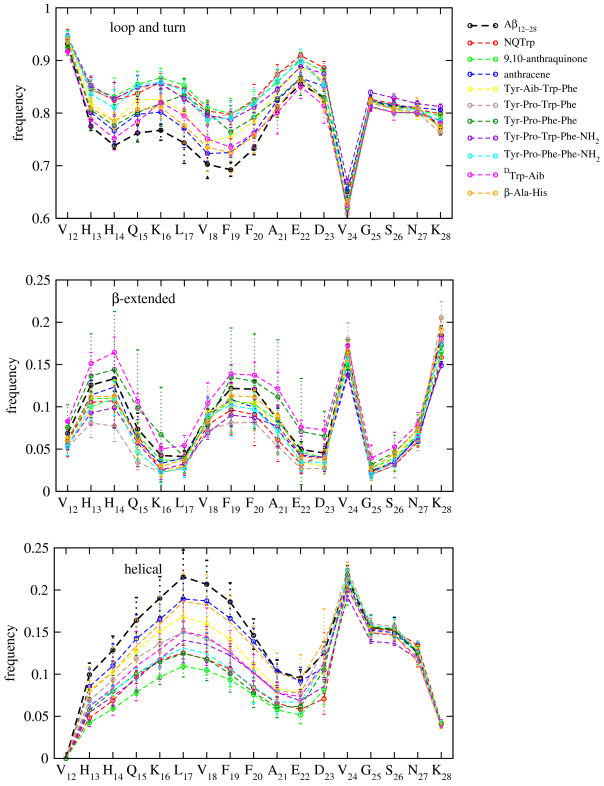
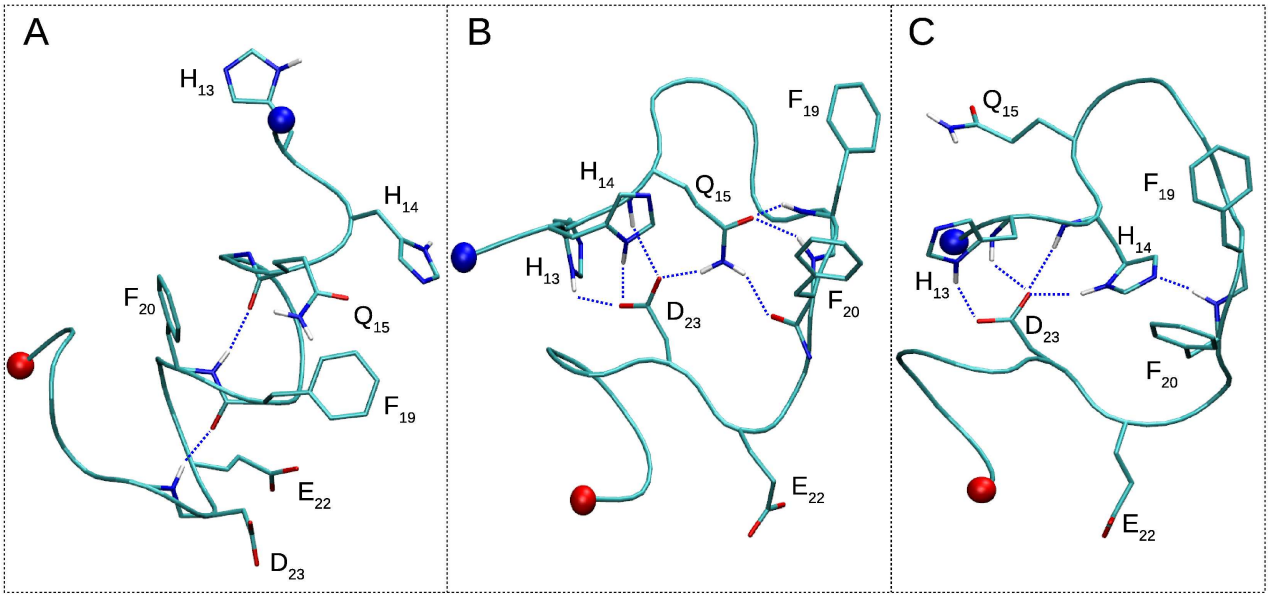


Figure 7



DISORDERED BINDING OF SMALL MOLECULES TO Aβ_{12-28}

Marino Convertino, Andreas Vitalis*, and Amedeo Caflisch*

Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

Supplemental Data

Address correspondence to:

Prof. Dr. Amedeo Caflisch, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland.
Tel: +41 44 635 55 21; fax: +41 44 635 68 62; E-mail: caflisch@bioc.uzh.ch.

Dr. Andreas Vitalis, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland.
Tel: +41 44 635 55 97; fax: +41 44 635 68 62; E-mail: a.vitalis@bioc.uzh.ch.

Table S1. Percentage of molecular dynamics snapshots similar to the representative loop conformation obtained in the presence of NQTrp using 1.5 Å and 2.0 Å cutoffs in RMSD computed over the Cα atoms of residues 14-24.

Inhibitor [a] **(cutoff 1.5 Å)**	**3 blocks** [b]				**6 blocks** [b]			
	μ	**σ**	**MAX**	**min**	**μ**	**σ**	**MAX**	**min**
Aβ_{12-28}	1.8	1.1	2.9	1.0	1.8	1.4	4.2	0.2
NQTrp	19.0	15.3	40.1	4.4	19	17.6	48.7	0.4
9,10-anthraquinone	8.2	6.3	15.5	0.2	8.2	10.8	30.4	0.2
anthracene	5.4	3.6	9.7	0.9	5.4	4.8	11.4	0.3
Tyr-Aib-Trp-Phe	4.5	1.6	6.3	2.4	4.5	4.9	12.5	0.1
Tyr-Pro-Trp-Phe	5.1	2.7	8.9	2.5	5.1	6.3	17.7	0.1
Tyr-Pro-Phe-Phe	5.2	2.4	8.3	2.4	5.2	4.8	14.6	0.2
Tyr-Pro-Trp-Phe-NH_2	7.4	3.8	10.9	3.4	7.4	5.8	15.3	0.4
Tyr-Pro-Phe-Phe-NH_2	4.3	2.8	7.4	2.1	4.3	3.3	8.0	0.2
^{D}Trp-Aib	6.9	6.5	16.1	1.7	6.9	11.2	31.7	0.2
β-Ala-His	1.4	1.4	3.3	0.1	1.4	1.4	3.8	0.1
Inhibitor [a] **(cutoff 2.0 Å)**	**3 blocks** [b]				**6 blocks** [b]			
	μ	**σ**	**MAX**	**min**	**μ**	**σ**	**MAX**	**min**
Aβ_{12-28}	2.9	1.3	4.3	1.6	2.9	1.7	5.5	0.9
NQTrp	22.5	14.0	41.6	8.5	22.5	16.3	50.0	5.4
9,10-anthraquinone	10.3	7.1	18.3	1.1	10.3	11.4	33.3	0.7
anthracene	6.8	3.7	11.1	1.8	6.8	4.4	12.7	2.2
Tyr-Aib-Trp-Phe	6.1	1.5	8.2	4.6	6.1	4.4	14	1.5
Tyr-Pro-Trp-Phe	7.6	3.3	12.2	4.9	7.6	5.4	19.1	3.2
Tyr-Pro-Phe-Phe	16.6	1.1	18.2	15.4	16.6	2.0	20.6	14.3
Tyr-Pro-Trp-Phe-NH_2	11.7	2.2	14.1	9.8	11.7	5.7	18.0	1.6
Tyr-Pro-Phe-Phe-NH_2	9.9	5.9	16.5	5.1	9.9	8.0	25.3	1.6
^{D}Trp-Aib	8.3	6.9	18.0	2.4	8.3	11.4	33.3	0.7
β-Ala-His	2.2	1.6	4.4	0.6	2.2	1.7	5.5	0.4

[a]Aib: α-aminoisobutyric acid; NQTrp: 1,4-napthoquinon-2-yl-L-tryptophan; β-Ala-His: carnosine.

[b]Block averaging was performed by dividing the 15 μs of total sampling into three 5 μs blocks (left) and six 2.5 μs blocks (right). μ and σ denote arithmetic mean and standard deviation, respectively, while MAX and min are the respective maximum and minimum values for each set of blocks.

Table S2. Decomposition of effective binding energy into individual terms. All values are given in kcal/mol.

Inhibitor [a]	sampling [g]	vdW [b]	ELEC [c]	FCTPL [d]	FCTNP [e]	TOTAL [f]
NQTrp	*full*	-15.8	-12.9	17.7	-3.0	-13.9
	1st block	-15.1	-13.4	18.1	-2.9	-13.4
	2nd block	-18.4	-15.2	20.3	-3.3	-16.6
	3rd block	-13.7	-9.9	14.6	-2.8	-11.8
9,10-anthraquinone	*full*	-6.6	-2.9	5.4	-1.5	-5.6
	1st block	-8.1	-5.4	8.1	-1.7	-6.9
	2nd block	-6.9	-3.1	5.7	-1.5	-5.9
	3rd block	-4.9	-0.3	2.6	-1.4	-4.1
anthracene	*full*	-7.7	0.1	1.8	-1.7	-7.5
	1st block	-8.1	-1.3	3.2	-1.8	-7.9
	2nd block	-6.1	2.35	-0.6	-1.6	-5.9
	3rd block	-8.8	-0.77	2.9	-1.9	-4.1
Tyr-Aib-Trp-Phe	*full*	-9.6	-8.8	12.1	-2.8	-9.2
	1st block	-9.5	-9.5	12.6	-2.8	-9.2
	2nd block	-8.8	-6.9	10.2	-2.7	-8.3
	3rd block	-10.5	-10.1	13.4	-2.9	-10.1
Tyr-Pro-Trp-Phe	*full*	-10.3	-10.7	14.1	-2.9	-9.8
	1st block	-11.5	-13.3	16.8	-3.1	-11
	2nd block	-9.9	-7.8	10.9	-2.8	-8.5
	3rd block	-10.3	-11.2	14.5	-2.9	-9.9
Tyr-Pro-Phe-Phe	*full*	-16.3	-14.7	19.1	-3.6	-15.6
	1st block	-16.9	-16.4	20.6	-3.6	-16.3
	2nd block	-16.8	-14.7	19.1	-3.7	-15.9
	3rd block	-15.3	-13.1	17.4	-3.5	-14.5
Tyr-Pro-Trp-Phe-NH_2	*full*	-11.9	-5.9	10.0	-3.3	-11.1
	1st block	-11.9	-7.0	11.1	-3.3	-11.2
	2nd block	-10.6	-4.3	8.2	-3.1	-9.9
	3rd block	-13.0	-6.4	10.8	-3.4	-12.0

Tyr-Pro-Phe-Phe-NH_2	*full*	-10.5	-7.0	10.5	-3.0	-10.0
	1st block	-12.7	-10.1	13.9	-3.3	-12.1
	2nd block	-7.2	-2.9	5.8	-2.6	-7.0
	3rd block	-11.2	-7.5	11.2	-3.1	-10.6
^{D}Trp-Aib	*full*	-4.5	-5.5	7.3	-1.8	-4.5
	1st block	-3.1	-5.5	6.8	-1.6	-3.3
	2nd block	-3.7	-2.9	4.6	-1.7	-3.6
	3rd block	-6.8	-8.1	10.3	-2.0	-6.6
β-Ala-His	*full*	-0.9	-4.9	5.7	-0.8	-0.8
	1st block	-0.5	-4.6	5.3	-0.8	-0.7
	2nd block	-0.5	-4.7	5.7	-0.7	-0.3
	3rd block	-1.7	-5.2	6.2	-0.9	-1.5

[a]Aib: α-aminoisobutyric acid; NQTrp:1,4-napthoquinon-2-yl-L-tryptophan; β-Ala-His: carnosine.

[b]vdW = van der Waals energy change upon binding

[c]ELEC = low-dielectric Coulombic energy upon binding

[d]FCTPL = change of electrostatic polarization upon binding calculated by FACTS (Haberthür and Caflisch, *J Comput Chem*, 2008)

[e]FCTNP = change of non-polar solvation energy calculated by FACTS using a surface tension-like parameter γ= 0.0075 kcal mol^{-1} $Å^{-2}$

[f]TOTAL = the total binding energy is the sum of the four preceding terms

[g]Block averaging was performed by dividing the 15 μs sampling into three 5 μs blocks. For further details, refer to Table 1 in the main text.

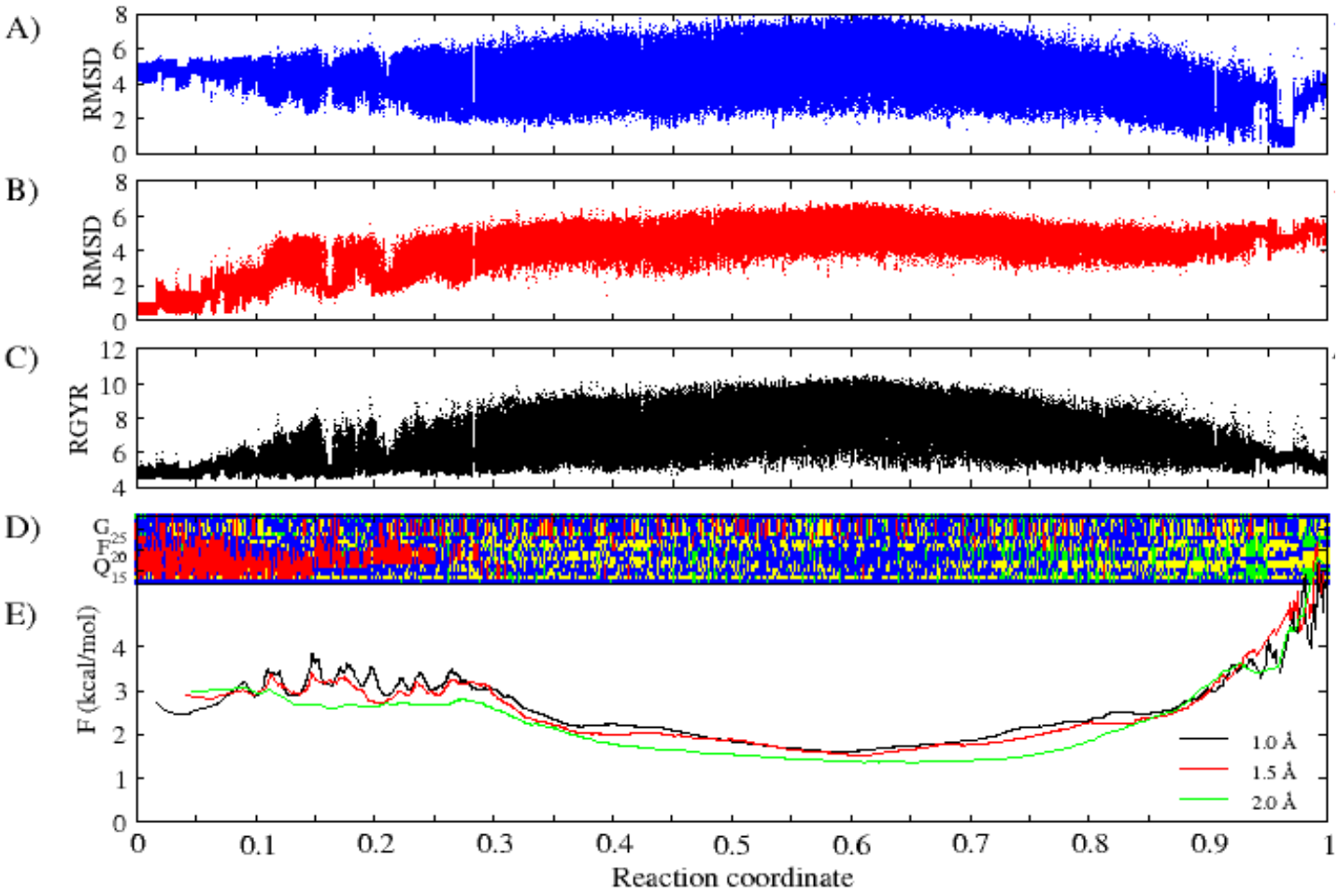


Figure S1. cFEP and other analyses of monomeric A$\beta_{12\text{-}28}$. The analysis in **A-D)** uses the snapshots rearranged according to the reaction coordinate in **E)**. **A)** Root mean square deviation (RMSD) of monomeric A$\beta_{12\text{-}28}$ from the cluster representative of the most populated node of A$\beta_{12\text{-}28}$ in presence of NQTrp. **B)** RMSD of monomeric A$\beta_{12\text{-}28}$ from its most populated conformer. **C)** Radius of gyration (RGYR) of monomeric A$\beta_{12\text{-}28}$. Cα atoms of residues 14-24 were used in RMSD and RGYR analyses. **D)** DSSP analysis of monomeric A$\beta_{12\text{-}28}$ (the color code is: helix in red; β-extended in green; loop and turn in blue; bend is plotted in yellow for clarity). The secondary structure analysis was calculated using the DSSPcont algorithm as implemented in Wordom (Seeber *et al*., *J Comput Chem* **32**, 1183 (2011)). **E)** cFEP plots starting from the most populated node of the conformational network as reference. Nodes are sorted according to their mean first passage time (mfpt) from the reference structure (for further details see Krivov and Karplus, *J \Phys Chem B*, **110**, 12689 (2006)). RMSD clustering over the Cα atoms of residues 14-24 of monomeric A$\beta_{12\text{-}28}$ with 1.0 Å (black), 1.5 Å (red), and 2.0 Å (green) cutoffs was used for cFEP analysis.

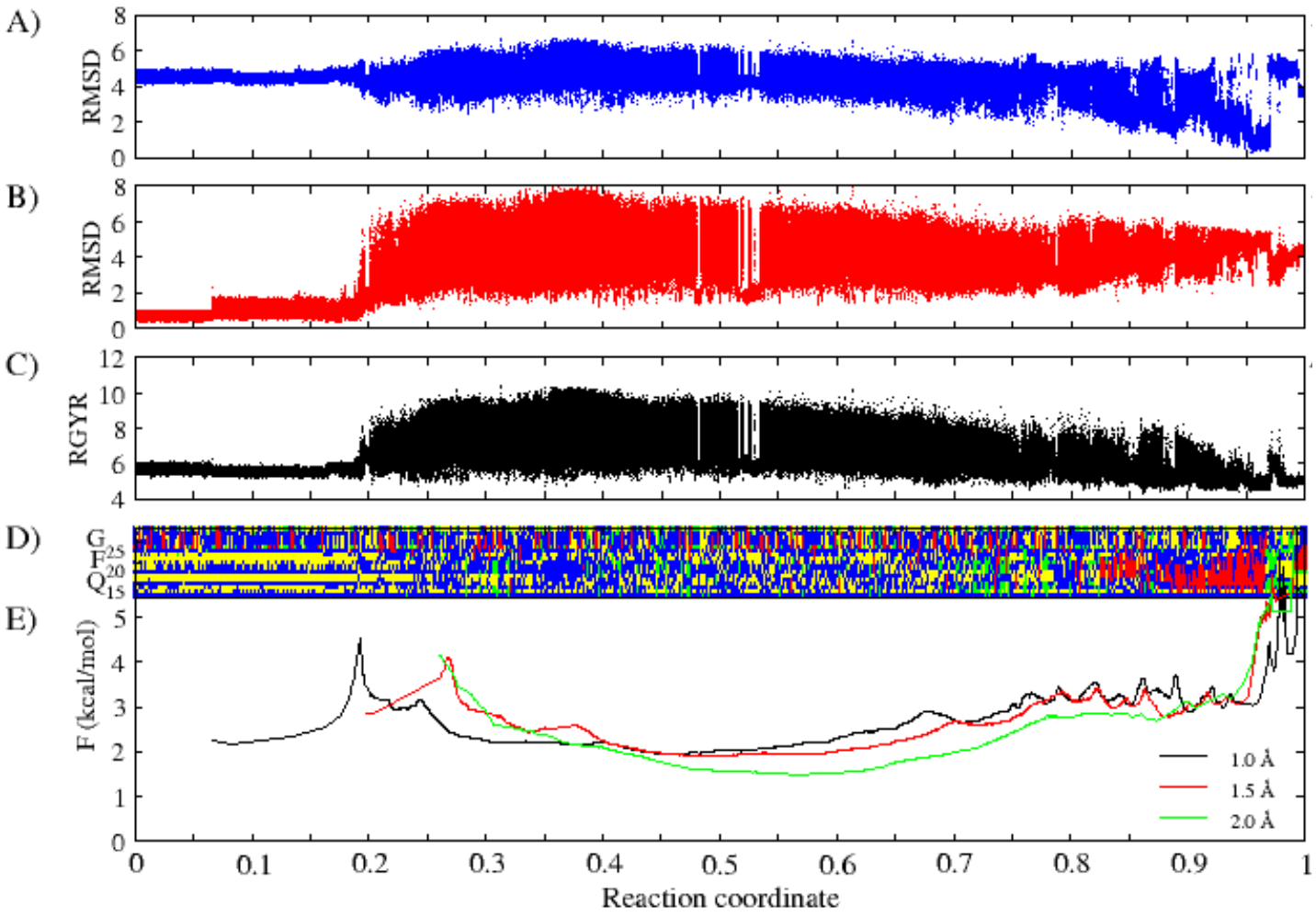


Figure S2. cFEP and other analyses of monomeric A$\beta_{12\text{-}28}$ in presence of NQTrp. The analysis in **A-D)** uses the snapshots rearranged according to the reaction coordinate in **E)**. **A)** RMSD of A$\beta_{12\text{-}28}$ in presence of NQTrp from the cluster representative of the most populated node of monomeric A$\beta_{12\text{-}28}$ alone. **B)** RMSD of A$\beta_{12\text{-}28}$ in presence of NQTrp from its most populated conformer. **C)** Radius of gyration (RGYR) of A$\beta_{12\text{-}28}$ in presence of NQTrp. Cα atoms of residues 14-24 were used in RMSD and RGYR analyses. **D)** DSSP analysis of monomeric A$\beta_{12\text{-}28}$ in presence of NQTrp (the color code is: helix in red; β-extended in green; loop and turn in blue; bend is plotted in yellow for clarity). **E)** cFEP plots starting from the most populated node of the conformational network as reference. Nodes are sorted according to their mean first passage time (mfpt) from the reference structure. RMSD clustering over the Cα atoms of residues 14-24 of monomeric A$\beta_{12\text{-}28}$ in presence of NQTrp with 1.0 Å (black), 1.5 Å (red), and 2.0 Å (green) cutoffs was used for cFEP analysis.

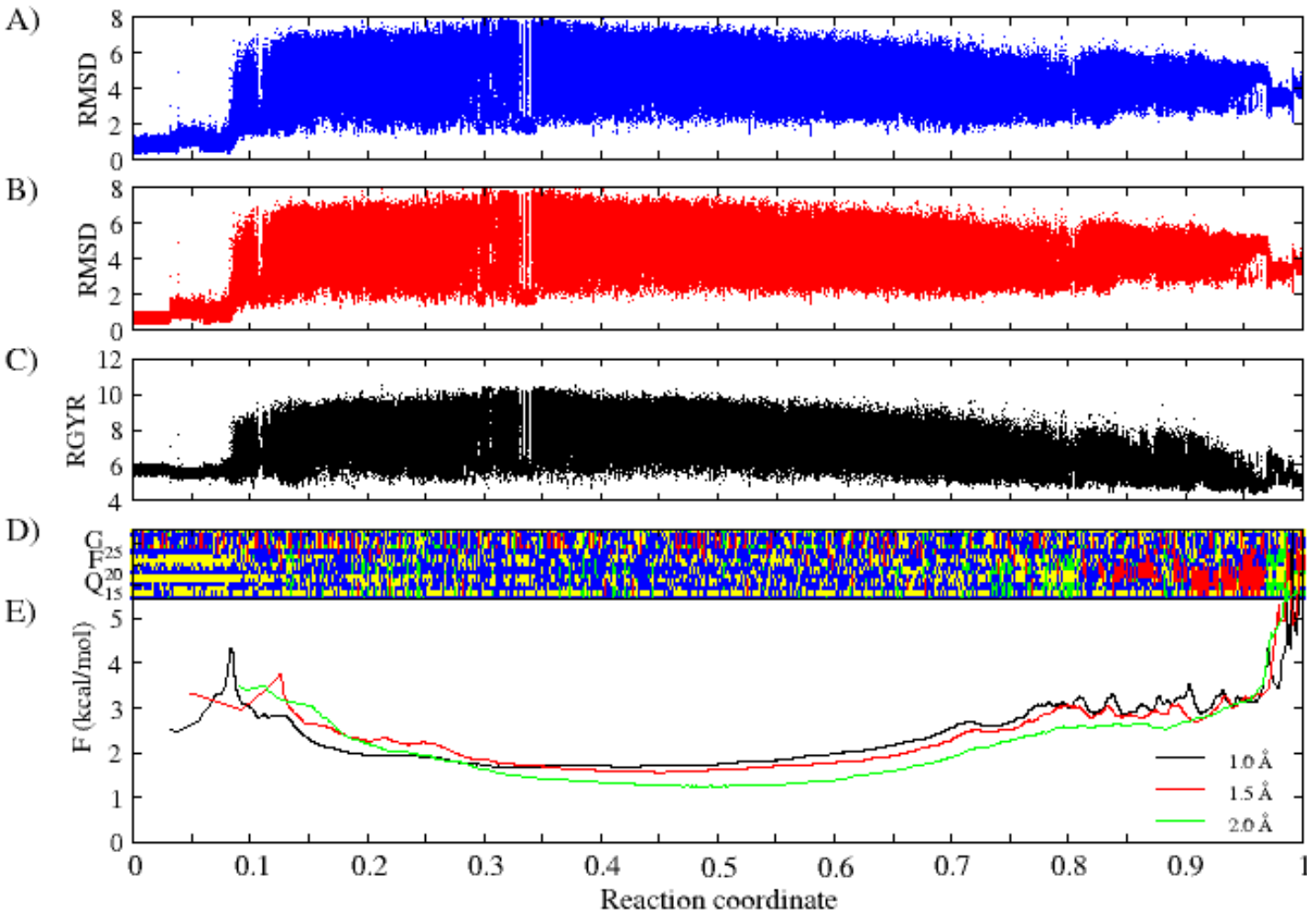


Figure S3. cFEP and other analyses of monomeric A$\beta_{12\text{-}28}$ in presence of 9,10-anthraquinone. The analysis in **A-D)** uses the snapshots rearranged according to the reaction coordinate in **E)**. **A)** RMSD of A$\beta_{12\text{-}28}$ in presence of 9,10-anthraquinone from the cluster representative of the most populated node of monomeric A$\beta_{12\text{-}28}$ in presence of NQTrp. **B)** RMSD of A$\beta_{12\text{-}28}$ in presence of 9,10-anthraquinone from its most populated conformer. **C)** Radius of gyration (RGYR) of A$\beta_{12\text{-}28}$ in presence of 9,10-anthraquinone. Cα atoms of residues 14-24 were used in RMSD and RGYR analyses. **D)** DSSP analysis of monomeric A$\beta_{12\text{-}28}$ in presence of 9,10-anthraquinone (the color code is: helix in red; β-extended in green; loop and turn in blue; bend is plotted in yellow for clarity). **E)** cFEP plots starting from the most populated node of the conformational network as reference. Nodes are sorted according to their mean first passage time (mfpt) from the reference structure. RMSD clustering over the Cα atoms of residues 14-24 of monomeric A$\beta_{12\text{-}28}$ in presence of 9,10-anthraquinone with 1.0 Å (black), 1.5 Å (red), and 2.0 Å (green) cutoffs was used for cFEP analysis.

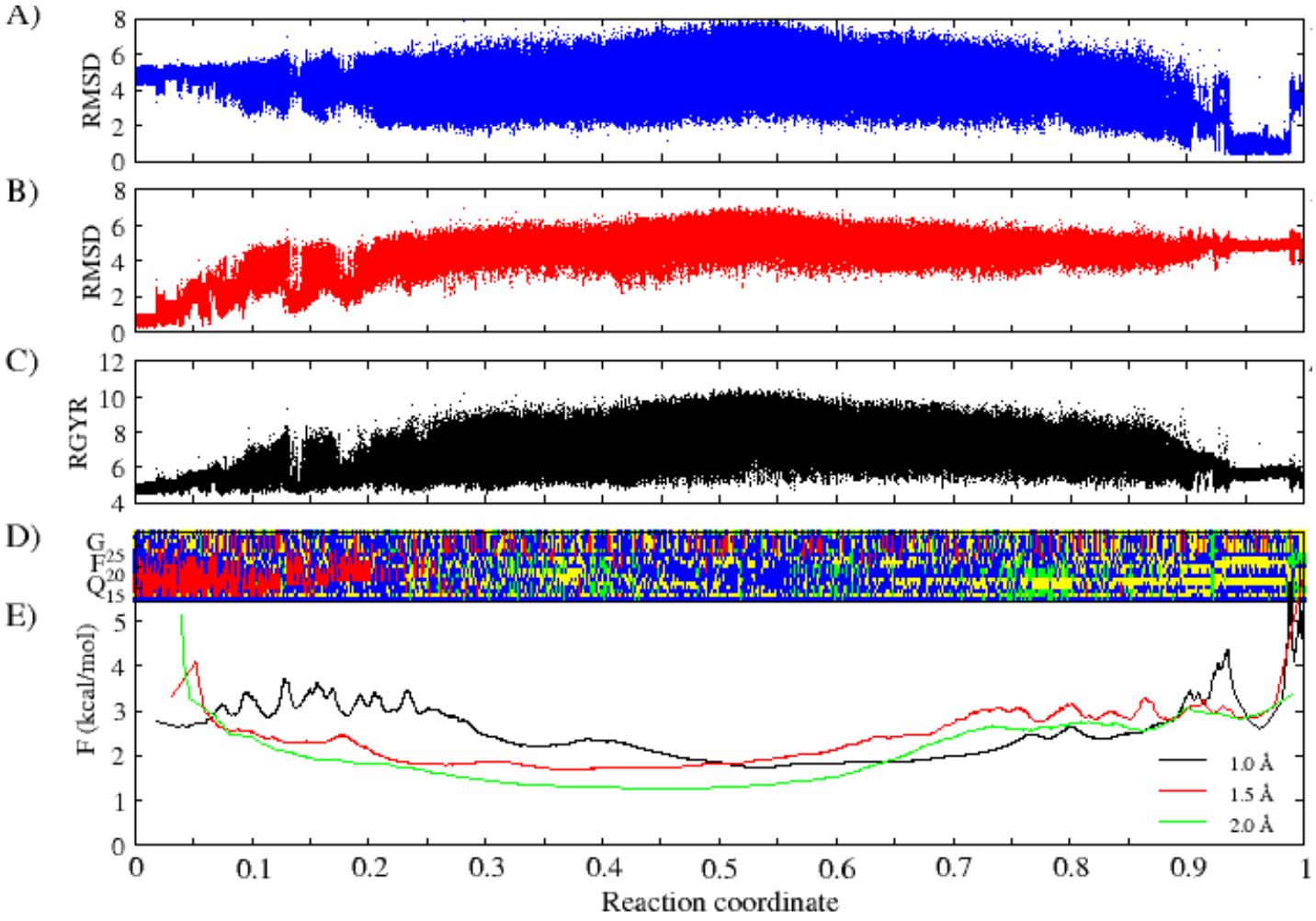


Figure S4. cFEP and other analyses of monomeric A$\beta_{12\text{-}28}$ in presence of anthracene.

Same as Figure S3 for anthracene.

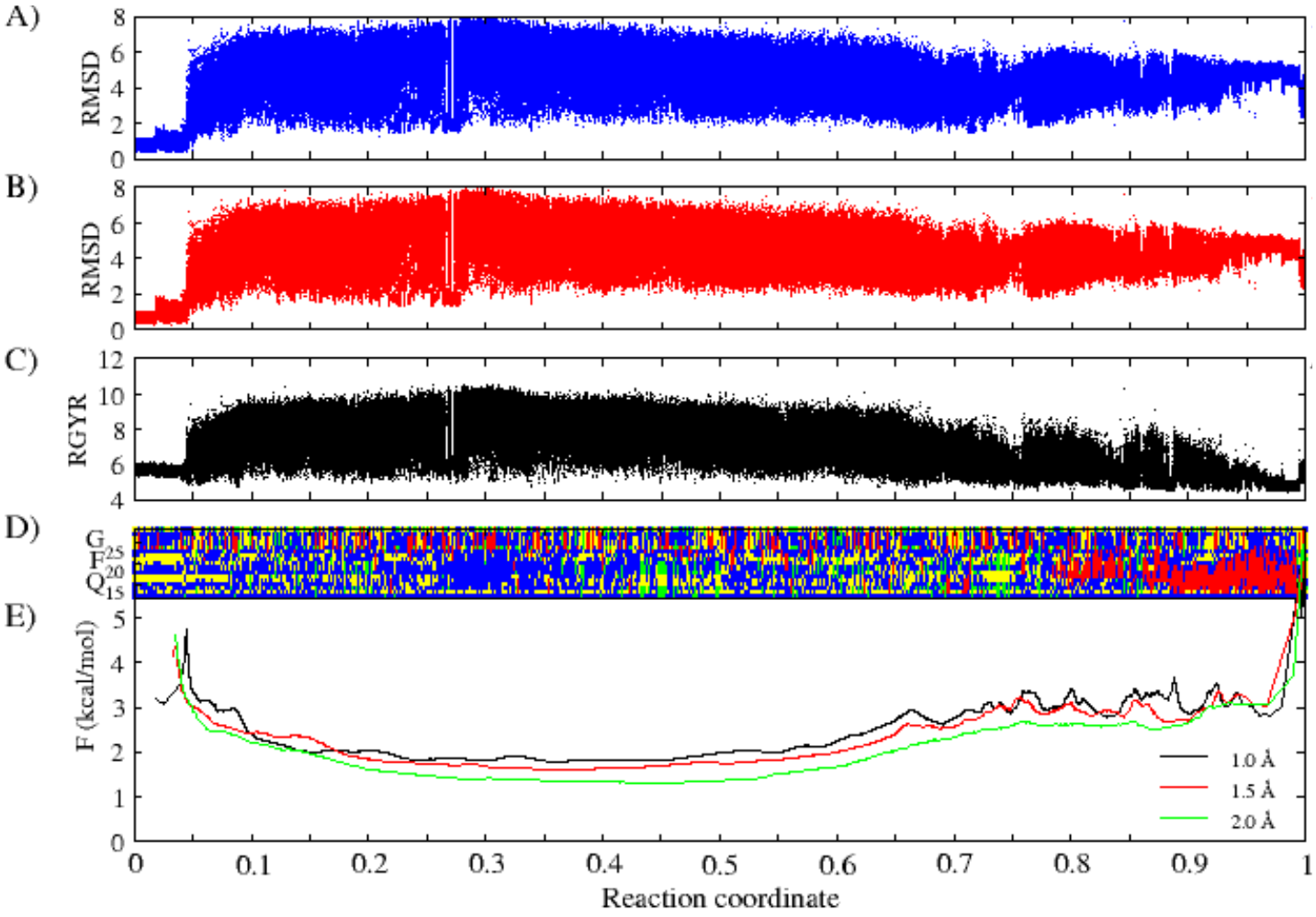


Figura S5. cFEP and other analyses of monomeric Aβ_{12-28} in presence of Tyr-Aib-Trp-Phe.

Same as Figure S3 for Tyr-Aib-Trp-Phe.

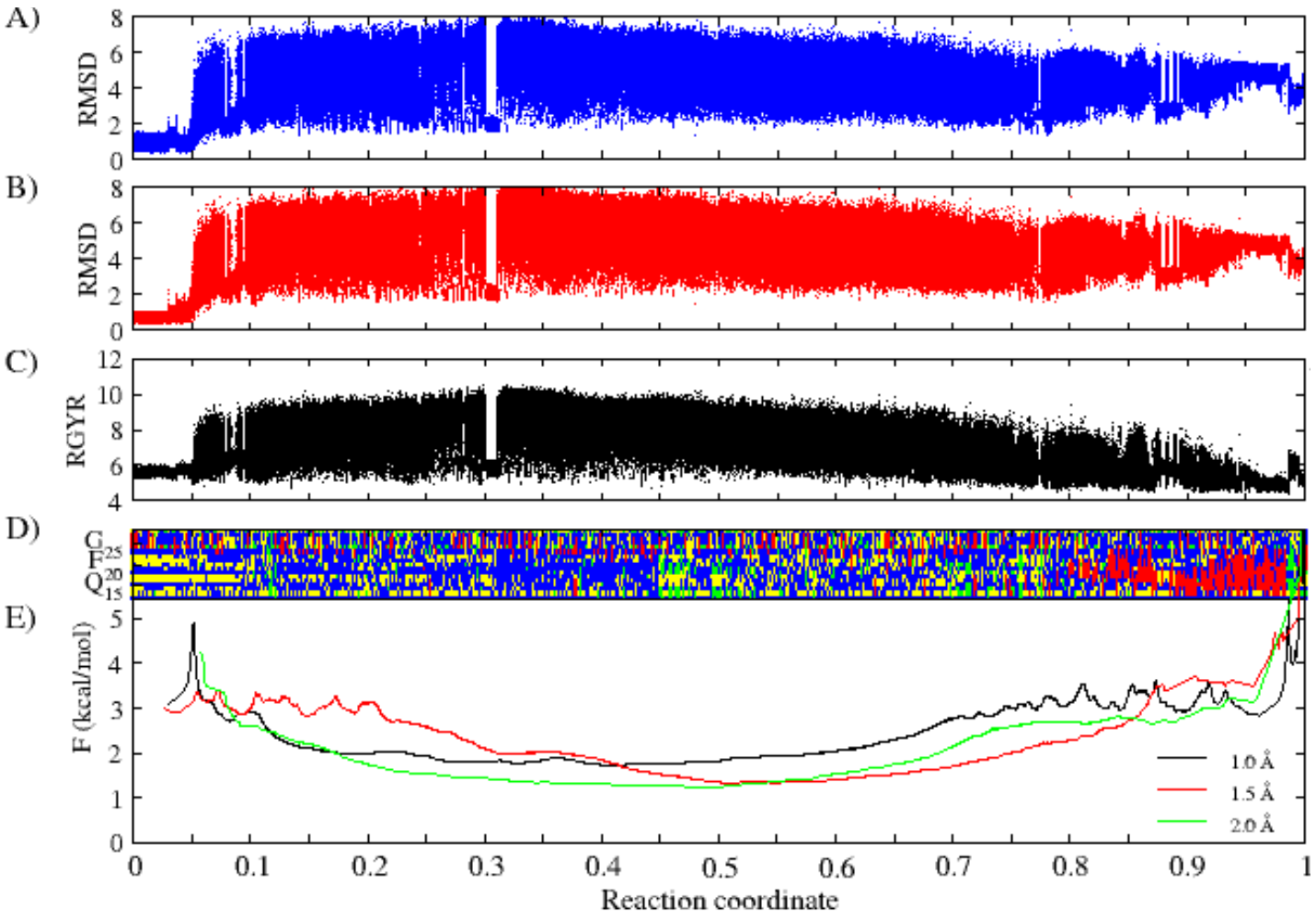


Figure S6. cFEP and other analyses of monomeric Aβ_{12-28} in presence of Try-Pro-Trp-Phe.

Same as Figure S3 for Tyr-Pro-Trp-Phe.

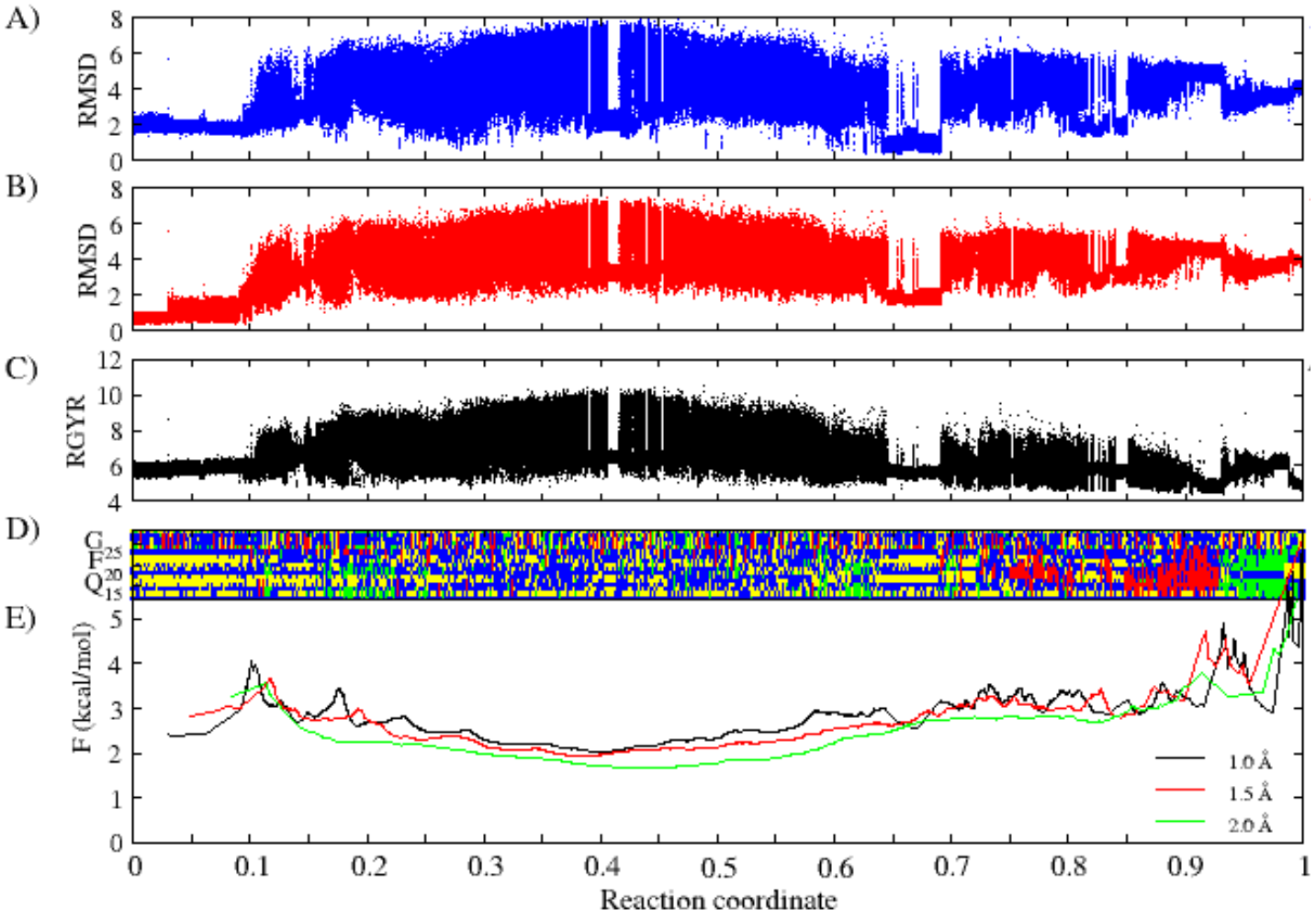


Figure S7. cFEP and other analyses of monomeric A$\beta_{12\text{-}28}$ in presence of Tyr-Pro-Phe-Phe.

Same as Figure S3 for Tyr-Pro-Phe-Phe.

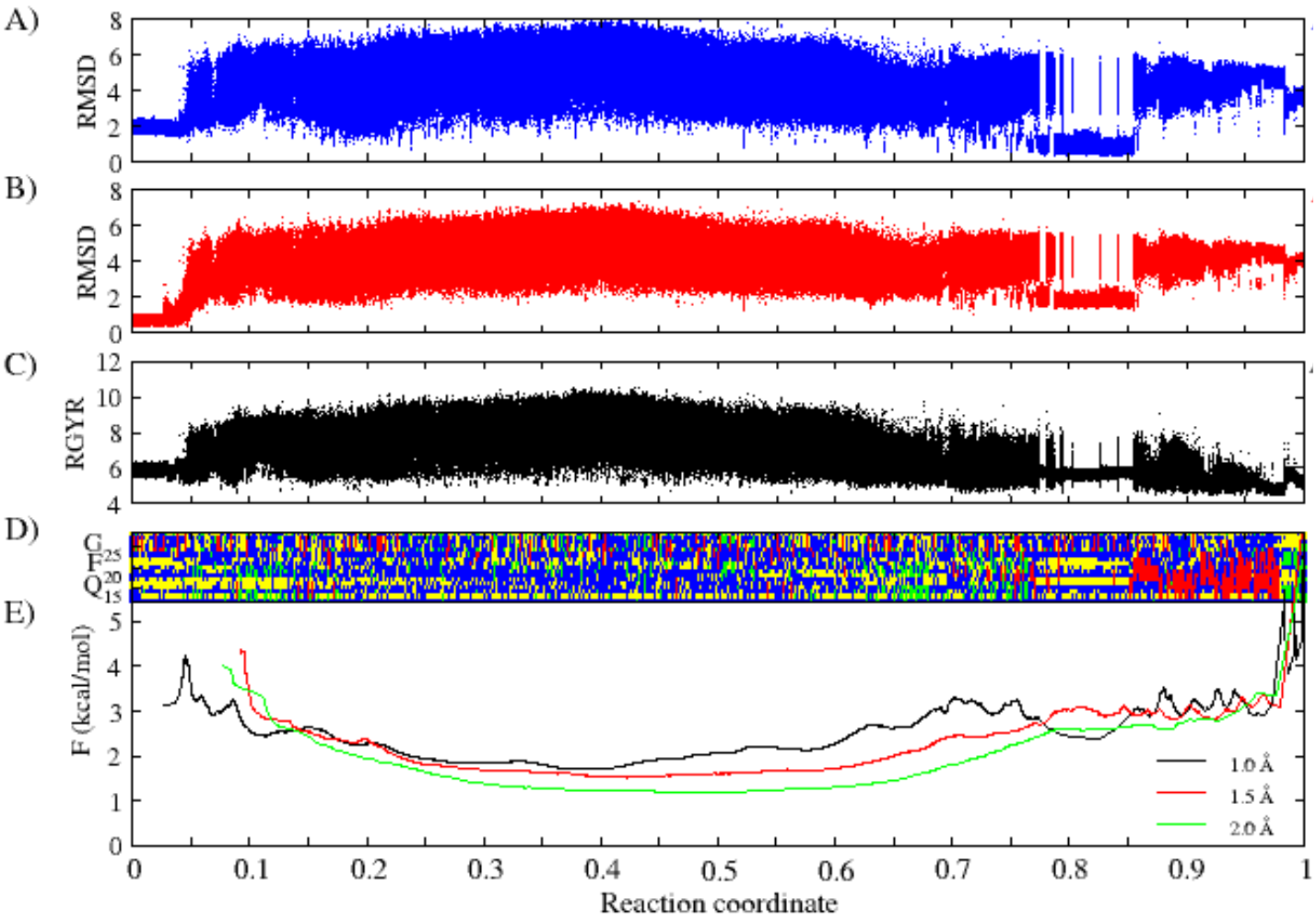


Figure S8. cFEP and other analyses of monomeric A$\beta_{12\text{-}28}$ in presence of Tyr-Pro-Trp-Phe-NH_2.

Same as Figure S3 for Tyr-Pro-Trp-Phe-NH_2.

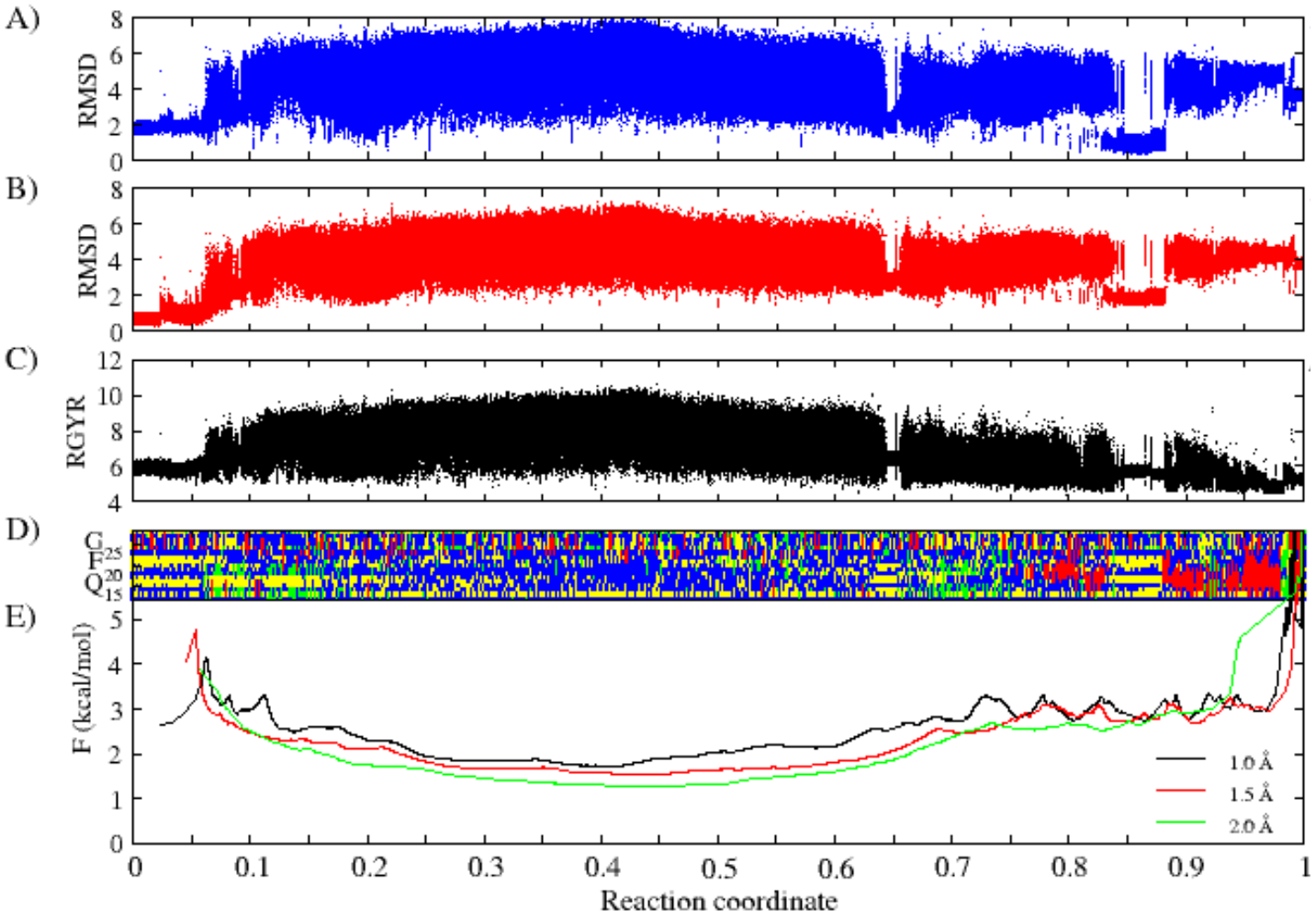


Figure S9. cFEP and other analyses of monomeric A$\beta_{12\text{-}28}$ in presence of Tyr-Pro-Phe-Phe-NH_2.

Same as Figure S3 for Tyr-Pro-Phe-Phe-NH_2.

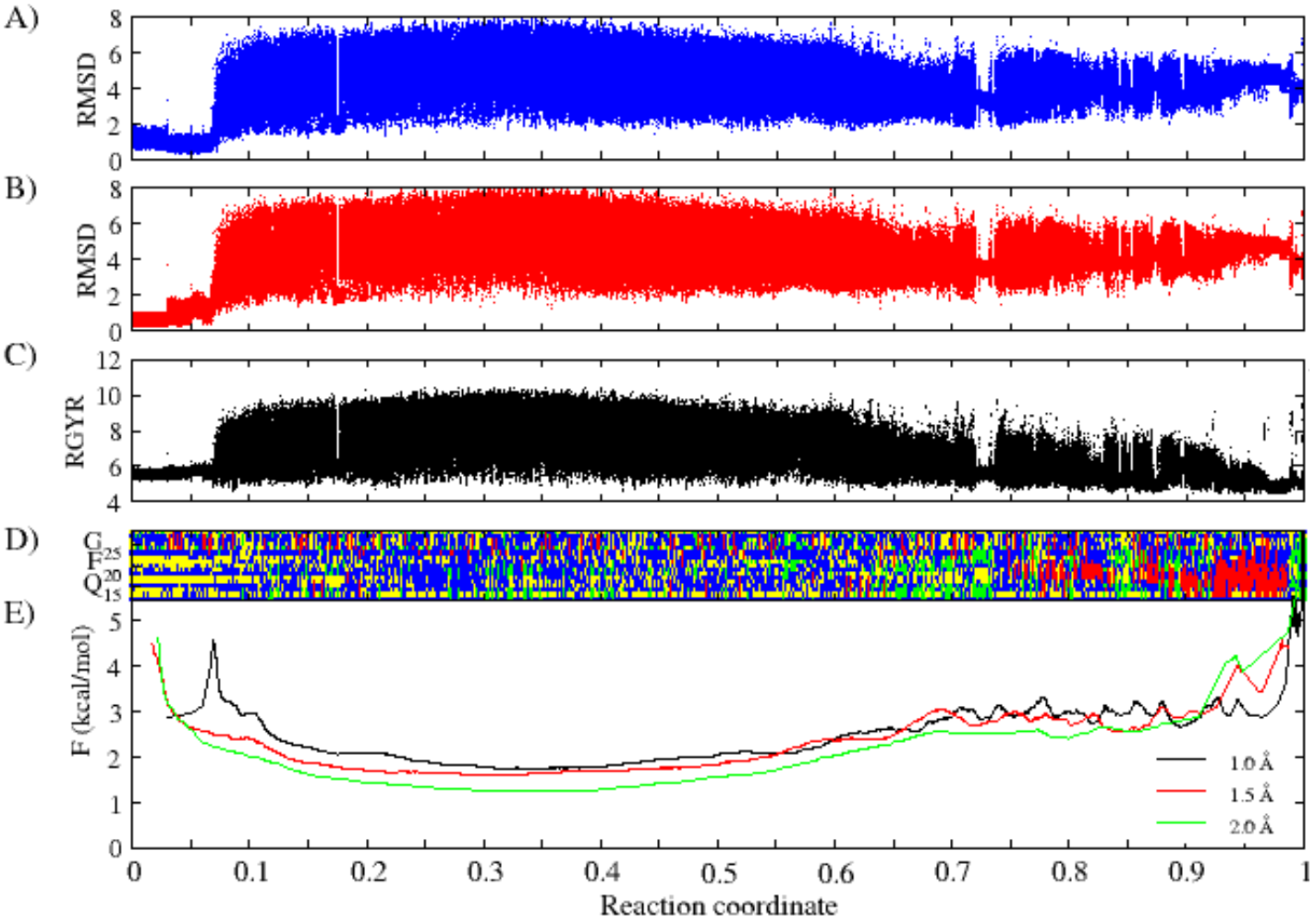


Figure S10. cFEP and other analyses of monomeric A$\beta_{12\text{-}28}$ in presence of ^{D}Trp-Aib.

Same as Figure S3 for ^{D}Trp-Aib.

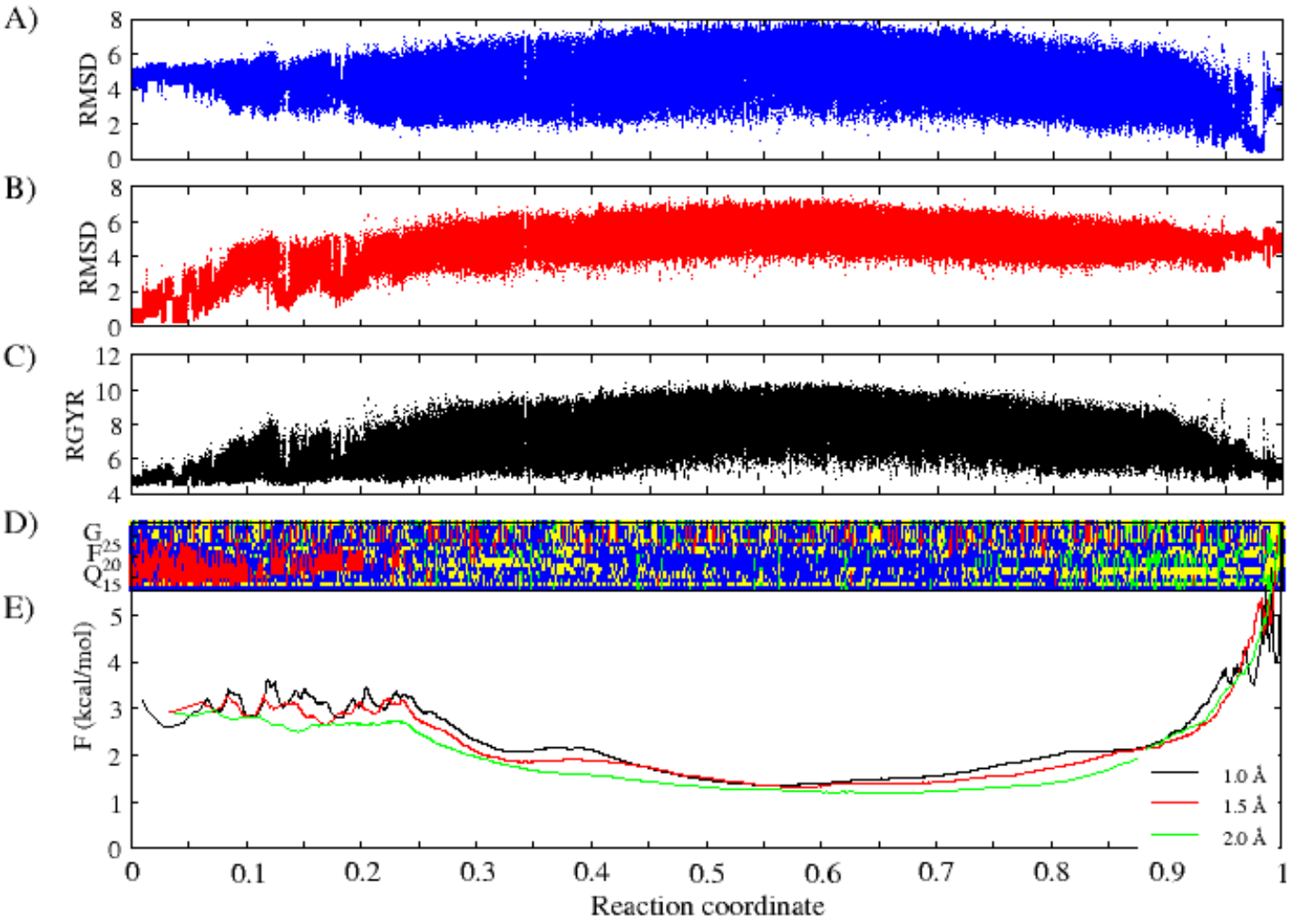


Figure S11. cFEP and other analyses of monomeric Aβ_{12-28} in presence of β-Ala-His.

Same as Figure S3 for β-Ala-His.

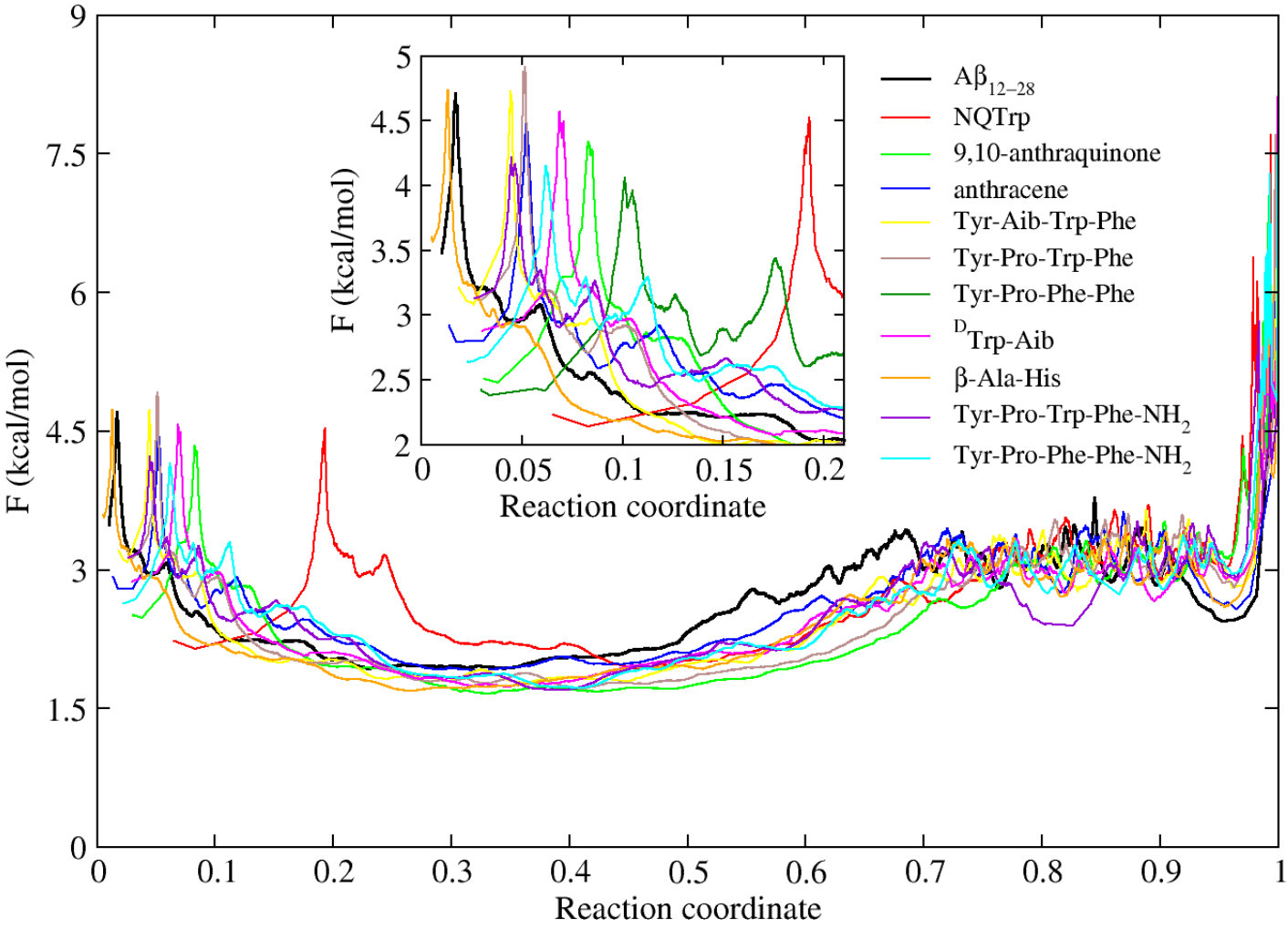


Figure S12. cFEP plots with the loop conformer as encountered in each simulation as reference state.

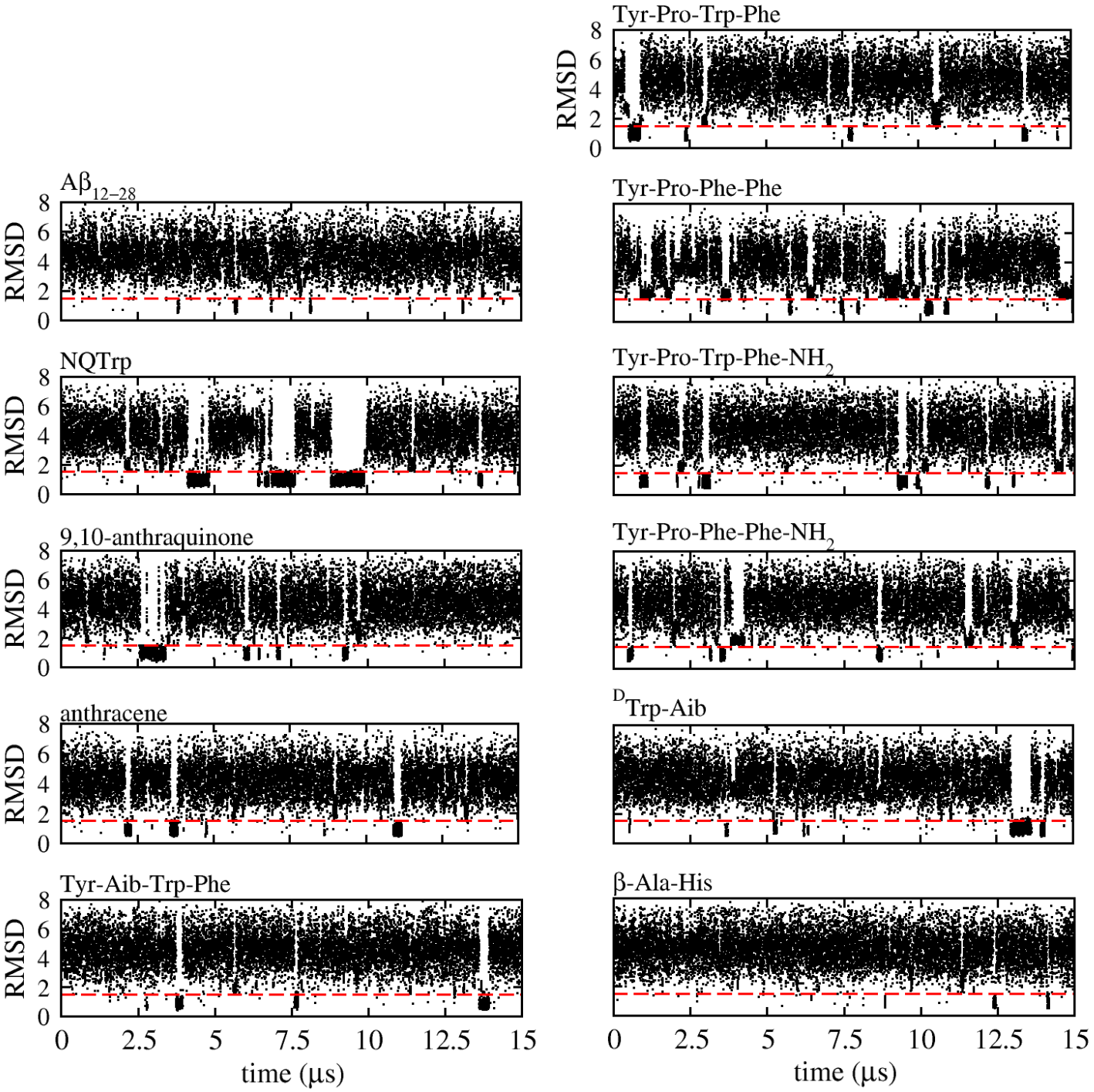


Figure S13. Time series of Cα RMSD from the representative node of the most populated cluster in simulations of Aβ_{12-28} in the presence of NQTrp (loop conformer). The Cα atoms of residues 14-24 were used to calculate the RMSD. The red dashed line at 1.5 Å is only a guide to the eye.

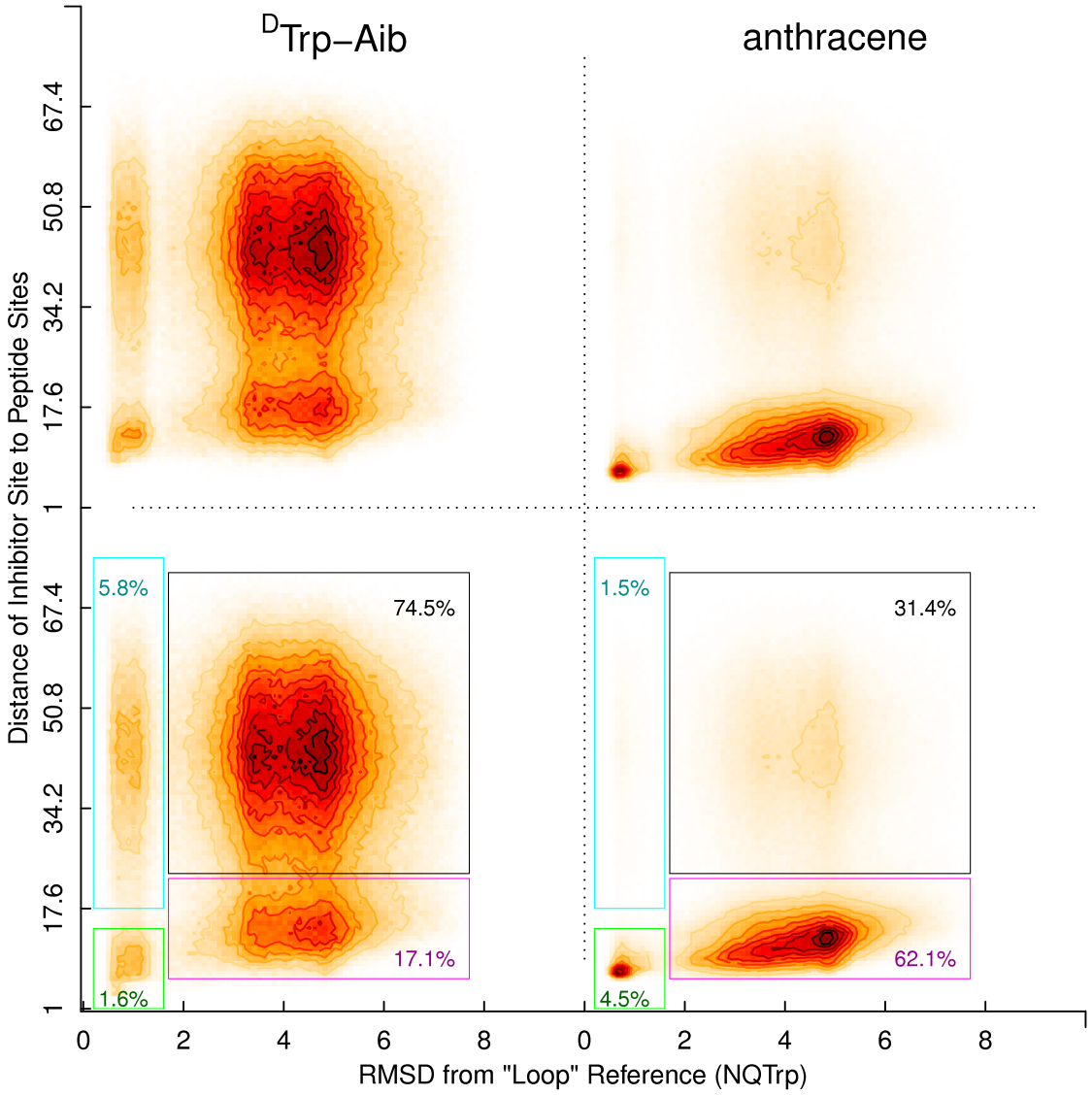


Figure S14. Two-dimensional histograms of the RMSD over Cα atoms of residues 14-24 to a reference structure and site-site distances between inhibitor and peptide. The chosen reference structure was the representative node of the most populated cluster for the system in which the loop structure was most strongly populated, i.e., A$\beta_{12\text{-}28}$ in the presence of NQTrp. Here, two different inhibitors are analyzed: ^{D}Trp-Aib (left column), and anthracene (right column). The sites on the peptide were the side chain oxygen atoms of Asp23 and Gln15. For ^{D}Trp-Aib, site-site correlation functions were computed to the polar indole nitrogen (lower left) or to the N-terminal, charged nitrogen (upper left). For anthracene, the sites considered were the four terminal carbons (lower right), and the two central carbons (upper right). In all cases, the distance axis can be used to distinguish bound from unbound states, and the RMSD axis can be used to distringuish loop-like from other conformations. Due to the good separation between states, the bottom row contains approximate sums of probabilities in the indicated, rectangular regions. These data indicate that anthracene

stabilizes the loop largely by direct interaction, whereas the impact of ^{D}Trp-Aib can also be indirect. The latter could be the result of both allosteric and memory effects.